OCTOBER
1 9 5 1
VOLUME 15
NUMBER 5

# JOURNAL OF CONSULTING PSYCHOLOGY

AMERICAN PSYCHOLOGICAL ASSOCIATION

# JOURNAL OF CONSULTING PSYCHOLOGY

EDITED BY LAURANCE F. SHAFFER

TEACHERS COLLEGE, COLUMBIA UNIVERSITY

ASSOCIATE EDITORS: EDGAR A. DOLL, *Devereux Schools, Devon, Pa.* WILLIAM A. HUNT, *Northwestern University.* E. LOWELL KELLY, *University of Michigan,* MORRIS KRUGMAN, *Board of Education, New York, N. Y.* BERTHA M. LUCKEY, *Board of Education, Cleveland, Ohio.* FRED MCKINNEY, *University of Missouri.* ANNE ROE, *New York, N. Y.* CARL R. ROGERS, *University of Chicago.* R. NEVITT SANFORD, *University of California.*

MARGARET K. HARLOW, *Managing Editor*

THE JOURNAL OF CONSULTING PSYCHOLOGY is the clinical journal of the American Psychological Association. It is devoted primarily to the publication of original research relevant to psychological diagnosis, psychotherapy and counseling, personality, and the dynamics of behavior. Case studies, theoretical contributions, descriptions of clinical techniques, and discussions of the professional problems of clinical psychology also appear in its pages.

## EDITORIAL OFFICE

Manuscripts submitted for publication, books and tests for review, and correspondence on editorial matters should be sent to the Editor, Laurance F. Shaffer, 525 West 120th St., New York 27, N. Y.

***Information for contributors:*** The important characteristics of a good article are the significance of its subject matter, the demonstrated soundness of its conclusions, its clarity, and its brevity. Clarity is improved by careful outlining, discriminating use of headings, and a simple, straightforward style of writing. For brevity, authors should strive to convey every thought without waste of words, and should avoid repetitions of methods, data or findings.

Some of the mechanical requirements for manuscripts are: Type manuscripts with wide margins, on white bond paper. Double space between *each and every line,* including references, footnotes, matter to be set in smaller type, and tables. Use only two types of headings: centered headings, and underlined paragraph headings. Type each table, however small, on a separate sheet, with Arabic number and title. Draw all figures in India ink in final form for reproduction, with draftsman's lettering, but type captions for figures on separate sheets. The *References* must follow the exact form used in APA journals. For further information on the details of preparing manuscripts, see:

> ANDERSON, J. E., AND VALENTINE, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.,* 1944, 14, 345-376.

There is a lag of about nine to twelve months between the acceptance of an article and its publication. Authors may secure immediate publication, however, by arranging for articles to be printed as extra pages, increasing the number of pages that subscribers receive. The cost to the author of such early publication is approximately $14.00 per printed page.

Authors may order reprints when returning proof. No reprints are furnished gratis for articles submitted after January 1, 1952. Authors are charged one-half of the extra cost of tabular material and cuts, and the full cost of changes made on the proof.

## BUSINESS OFFICE

Communications concerning subscriptions, change of address, claims for the nonreceipt of a number, advertising, and other business matters should be sent to the American Psychological Association, Inc., 1515 Massachusetts Ave., N. W., Washington 5, D. C.

The Journal of Consulting Psychology is published bi-monthly in February, April, June, August, October, and December, the yearly volume comprising approximately 480 pages. The subscription per year is $5.00, foreign $5.50, single number $1.00.

Copyright 1951 by the American Psychological Association, Inc. Printed and mailed by the Dentan Printing Co., 23 West Colorado Ave., Colorado Springs, Colo., and entered as second-class matter at the Post Office at Colorado Springs, Colo. Acceptance for mailing at special rate of postage provided for in section 538, Act of February 25, 1925, authorized May 27, 1950.

JOURNAL OF CONSULTING PSYCHOLOGY

October, 1951 Vol. 15, No. 5

## CONTENTS

The Problem of Anxiety in Client-Centered Therapy: Ernst G. Beier - - - - - 359

Evaluation of Group Psychotherapy by Means of Performance Tests: Henry N. Peters and Francis D. Jones - - - - - - - - - - - - - - - 363

Predicting Continuation in Therapy by Means of the Rorschach Test: Lawrence S. Rogers, Joanne Knauss, and Kenneth R. Hammond - - - 368

The Generality of Psychiatric Syndromes: J. R. Wittenborn and J. D. Holzberg - 372

Equivalent Test and Mental Ages for the WISC: David Wechsler - - - - - - 381

Some Evidence on the Validity of the WISC: John J. Pastovic and George M. Guthrie - - - - - - - - - - - - - - 385

The Regression of Schizophrenic Language: Robert B. Ellsworth - - - - - - 387

Epileptic Intelligence: A. Louise Collins - - - - - - - - - - - - - - - 392

Induced Frustration and Punishment-Reward Expectancy in Thematic Apperception Stories: Vaughn J. Crandall - - - - - - - - - - - - - - 400

Projective Pictures as Interview Devices: H. K. Heppell and V. C. Raimy - - - 405

Some Effects of Color on Thematic Fantasy: Glen A. Brackbill - - - - - - - - 412

The Stimulus Values of the Szondi Pictures: E. Szollosi, D. E. Lamphiear, and H. L. Best - - - - - - - - - - - - - - - - - - 419

Psychological Case Reporting at Psychiatric Staff Conferences: Jules D. Holzberg, Salvatore L. Alessi, and Murray Wexler - - - - - 425

The Effect of Brief Stimulus Electroconvulsive Therapy Upon Psychological Test Performances: Isidor W. Scherer - - - - - - - - - - 430

Norms, Reliability, and Forms of the MMPI: A. R. Gilliland and Russell Colgin 435

An Original Test of Symbol Arrangement Validated on Organic Psychotics: Theodore C. Kahn - - - - - - - - - - - - - - - - - - - - - - 439

New Books and Tests - - - - - - - - - - - - - - - - - - - - - - - - 445

# THE PROBLEM OF ANXIETY IN CLIENT-CENTERED THERAPY

ERNST G. BEIER
SYRACUSE UNIVERSITY

THE concepts systemized by Rogers [4] into a therapeutic philosophy need more careful scrutiny. Idioms, such as, "permissive atmosphere," "acceptance," "integrity of the individual," and "reflection of feeling," first used so definitely by him, have begun to lose both specificity and meaning. The analysts, the so-called "directivists," and the moralists claim that they too believe in acceptance, that they respect the integrity of the individual, that they believe in a growth principle, and desire to create a permissive atmosphere. It seems that such factors as acceptance, permissive atmosphere, and integrity of the individual can easily be integrated into differently oriented theories. A fair statement is that it would be difficult indeed for client-centered therapy to claim uniqueness with any one of these terms.

"Client-centeredness" is a good example. A major emphasis of Roger's contribution is expressed in his belief that the client can restore himself to healthier patterns of behavior if he is placed in an appropriate therapeutic climate. The therapist is advised not to assist the client directly, other than by clarifying the content and feeling of his statements. This is urged in an effort to differentiate client-centered therapy from interpretive therapies. It soon became evident, however, that reflections of feelings as well as reflections of content required judgments by the therapist. The therapist must choose which of many client's statements he should single out for reflection [2] and with what intensity he should reflect them. There is increasing recognition of the importance of the therapist even in the client-centered therapeutic situation [5]. It thus becomes difficult to claim that client-centeredness is a unique variable of client-centered therapy.

The present paper explores the idea that different orientations in psychotherapy differ principally in techniques for handling client anxiety. All therapeutic approaches recognize that dealing with anxiety is a central problem. Techniques, however, differ widely. It is in its technique for dealing with client anxiety that client-centered therapy may have established its uniqueness.

Client-centered therapy is foremost in its request that the therapist's aim should be one of anxiety-reduction. It is felt that in a nonthreatening situation, the individual's capacity for growth is freed [6]. Other orientations maintain that anxiety is not necessarily to be reduced in therapy. Part of the technique of some schools of psychoanalysis [1, 3] is the utilization of the nonconscious needs of the analysand in the therapeutic situation. Dependency anxiety (transference neurosis) and shock anxiety are familiar to the student of psychoanalysis. Anxiety aroused by depth interpretations, by analysis of the transference phenomenon, by "breaking down of unconscious defenses" is to many the very essence of the therapeutic relationship. Certainly moral (religious) and so-called directive or advice therapies utilize authority (anxiety generation) as one of their therapeutic tools [7]. Biological as well as psychological threats of disease, pain, guilty conscience, and abnormality, are utilized to keep the individual anxious for his own eventual happiness, as so brilliantly the Grand Inquisitor points out to Jesus (Dostoyevsky). It is clearly in the attempt to establish the nonthreatening anxiety-reducing situation that client-centered technique differs from other therapies.

An analysis of the counseling records of the Psychological Services Center of Syracuse

University reveals that one out of three clients voluntarily terminate counseling after less than six contacts. Generally, the counselors seem to agree that they "lost" these clients; i. e., that they did not seem to have gone as far as they "could" have gone; that something in the counseling situation went amiss.

On closer scrutiny, the clients who leave a psychotherapeutic situation "early" fall into four categories: (1) Clients who came because of slight situational disturbances and who are genuinely helped within a few hours; (2) clients who escape inexperienced counselors; (3) clients who have lost some symptoms and *think* they have been helped; and (4) clients who "give up" because the therapeutic situation is too demanding.

While the first two categories are self-explanatory, the last two categories concern us here. What went amiss with the "nonthreatening" situation? It seems that in the third category, symptom removal came about too quickly; i. e., the "nonthreatening" situation may have reduced anxiety to a point where no further help was desired. In the fourth category the situation may not at all have been perceived as "nonthreatening" by the client. It apparently *aroused* anxieties that could no longer be dealt with in the situation.

To understand these failures, we have to investigate the principal technique of client-centered therapy more carefully. The therapist communicates to the client only his understanding of such facets of the client's problems as the client himself has voiced and ideated, i. e., is aware of. Let us call these responses "aware-need reflections." It is argued here that as some clients enter the therapeutic situation with severe ambivalence toward facing their problems, "aware-need reflections" not only fail to communicate the therapist's deep understanding, but actually imply severe misunderstanding to the client. Some clients come to therapists because they are unable to voice and understand their problems, because they are unable to handle their own anxiety. They experience needs they cannot cope with, needs that manifest themselves only in vague anxiety and tenseness. Let us call such unexpressed needs "unaware needs."

The client-centered therapist reflects the client's aware needs with the hope that this procedure will establish the desired nonthreatening situation. The situation in turn permits the client to dig into the pool of unaware needs. The argument is that the less threatening the situation, the more likely will unaware needs be faced.

Some clients, however, seem to be unable to proceed in this manner. The therapist's reflection of their aware needs may not convey complete understanding to them. These clients are dimly aware of needs that they cannot voice and that apparently are not understood by the therapist. They come into the therapeutic hour hoping to establish dependency, receive advice, and find someone who understands them better than they understand themselves. Instead, they face the therapist who makes a point of *not* understanding them better than they do themselves. Negative feelings about therapy may be experienced (and that again is not reflected unless *expressed*) which prevents them from producing new material to talk about, and such clients sometimes get the feeling that they talk in circles and get nowhere. These individuals, because of a feeling that the hour makes overwhelming demands on them, quit in desperation (category 4).

Another effect may be achieved. In voicing his aware needs, and in the exclusive attention of the therapist to these needs, the client may be reinforced in his belief that he has gotten hold of his problem. On experiencing some symptom removal, he breaks off the contacts because he feels that he has been helped (category 3).

Response to unaware needs as here contemplated serves an important therapeutic function. Let us view the client entering the therapeutic situation with some ambivalence toward it. He *wishes* to enter, and at the same time *resists* entering a therapeutic relationship. His resistance to enter the relationship is based on a feeling of danger (anxiety). The danger is overwhelming. In responding to the vague unaware anxiety of the client, the therapist communicates to the client: "You can enter this situation safely; I can understand the danger you see even though you cannot express it yourself." Responding to an unaware need here serves to *support the client's own desire*

*to stay in the therapeutic situation.* The client feels deeply understood. It is recommended here that therapists attempt to recognize and accept clients' unaware needs, such as resistance, dependency feelings, and dominance, in situations that bring clients' ambivalences toward the relationship to the fore.

For the purpose of practical illustration a classification of specific techniques is added. The categorization of the therapist's responses is necessarily arbitary, but it is undertaken for the purpose of pointing out the problem.

Below the *Category* is listed first. Under *Process,* the technique will be described, and under *Communication,* there is a judgment of the meaning of the counselor's statement to the client. Under *Effect,* an evaluation of the statement on the mutual dependence relationship is given.

1. *Category*: "Reflection of Content"; *Process*: restatement of communication; *Communication*: "I understand you in your verbal communication"; *Effect*: loose dependency bond.

2. *Category*: "Reflection of Feeling, Expressed"; *Process*: restatement of the feelings reported; *Communication*: "I understand you in your intimate feelings, and don't fear, I am *not* trying to go 'deeper' than you are yourself"; *Effect*: loose but somewhat greater dependency on part of the client; he needs the counselor in order to be intimately understood.

3. *Category*: "Reflection of Feeling, Implied"; *Process*: restatement of the feelings implied toward the counselor or counseling situation ("you really resented me for saying this," or "this was difficult for you to say"); *Communication*: "I am trying to understand you fully here, as to what you are saying and as to what you are meaning, even though you didn't really express it"; *Effect*: dependency feelings—perhaps the first of techniques here termed "unaware-need reflection."

4. *Category*: "Area Reflection"; *Process*: statement pointing out the similarity among several feelings that have been observed during the hour; *Communication*: "I am trying to understand you in what you are saying and in what you are implying, and in addition, I accept your need to rely on somebody that seems to understand you better than you understand yourself"; *Effect*: strong dependency, strong reliance on judgment of therapist, another "unaware-need reflection."

5. *Category*: "Total Contact Reflection"; *Process*: statement pointing out the similarities among feelings that have been observed during the total contact; *Communication*: same as in "area reflection," only the acceptance of the client's need for dependency is perhaps more thorough here; *Effect*: same as in "area reflection."

6. *Category*: "Transference Reflection"; *Process*: statement pointing out the similarities of feelings that were expressed by the client toward other individuals *and* toward the counselor; *Communication*: "I am trying to understand what you are saying and what you are implying; I want to point out to you that I realize that you have feelings toward me which you describe having had for others. In addition, I temporarily accept the role you place me in, in order to receive help from somebody who understands your feelings better than you do yourself"; *Effect*: dependency relationship, "unaware-need reflection" creating increased transference, but also the key to working through transference.

It is apparent here that the term "reflection" is used and not the term "interpretation." Reflection here means that the therapist attempts to react to the client's feelings that have either been expressed directly or have been observed during his contact with the client. As stated earlier, these techniques are not thought of as applicable to all clients but only to those where the detour of a dependency relationship has to be taken into consideration and where the client's own nonexpressed desire to stay in therapy has to be supported.

The question arises whether or not the client-centered therapist can reflect "unaware needs" without destroying the basic tenets under which he operates. The basic tenets are: (1) growth occurs in freedom from threat, (2) in this freedom the primary responsibility for growth should be the client's. Do we violate these tenets if we reflect needs and anxieties that we observe but that have not been expressed by the client?

These criticisms are well taken with a majority of cases, but in consideration of failures in client-centered therapy a varying emphasis on the two basic tenets may be necessary. Some clients do not seem to be ready to feel understood by a therapist who conveys to them acceptance of what they are. They feel that they are more than they can *voice* at this moment, and they desire the support of the therapist for this unexpressed part of themselves. Not to get this support may be threatening to them. The support afforded these clients may make the difference. It may lead individuals otherwise not reached by client-

centered therapy, to the point where they can accept responsibility for their own psychological growth. In this manner many clients who do not stay with the therapist because of the therapist's too strict application of the two basic tenets will be able to enter the therapeutic situation, and client-centered therapy may serve more diversities of maladjustment.

*Received January 22, 1951.*

## References

1. Alexander, F., and French, T. M. *Psychoanalytic therapy.* New York: Ronald, 1946.
2. Collier, R. M. A basis for integration rather than fragmentation in psychotherapy. *J. consult. Psychol.*, 1950, **14**, 199–205.
3. Reich, W. *Charakter-analyse.* Vienna: Sexpol Verlag, 1933.
4. Rogers, C. R. *Counseling and psychotherapy.* Boston: Houghton Mifflin, 1942.
5. Seeman, J. A study of the process of nondirective therapy. *J. consult. Psychol.*, 1949, **13**, 157–168.
6. Snyder, W. U. An investigation of the nature of non-directive psychotherapy. *J. gen. Psychol.*, 1945, **33**, 193–223.
7. Thorne, F. C. Principles of directive counseling and psychotherapy. *Amer. Psychologist*, 1948, **3**, 160–165.

# EVALUATION OF GROUP PSYCHOTHERAPY BY MEANS OF PERFORMANCE TESTS[1]

HENRY N. PETERS AND FRANCIS D. JONES
VETERANS ADMINISTRATION HOSPITAL, NORTH LITTLE ROCK, ARKANSAS

THE Porteus Mazes and the Mirror-Tracing Test have been found in recent studies to have a certain amount of validity as measures of social adjustment. Porteus [3, 4] has reported reliable differences in qualitative maze score between various socially maladjusted groups, criminals and delinquents among others, and groups of normal subjects. In a later publication [5] it was shown that quantitative scores, and to some extent qualitative scores, on the mazes parallel levels of social adjustment in patients with prefrontal lobotomy. Peters [1] has demonstrated that the time taken in mirror-tracing is markedly longer on the average for criminals, delinquent adolescents, and psychotic patients than it is for groups of normal subjects.

Since the basic purpose of psychotherapy is to improve social adjustment, and since these two tests have been found to correlate with social adjustment, they were included in a battery which was given to patients before and after a period of group psychotherapy. The purpose was to determine whether or not these performance tests are sensitive to the improvement wrought by the therapy, assuming of course that the therapy did bring about some improvement. This assumption is supported informally by reports of changes in the patients' ward behavior, by doctor's progress notes, by privileges and discharges granted, as well as by observations of the patients' behavior in the group. It must be admitted, however, that in this study the question of what was being measured, group psychotherapy or test sensitivity, was somewhat ambiguous. If the results of such a study should be negative, there would still remain the problem of deciding which of two alternatives was true, namely, that the therapy was ineffective or that the tests were insensitive. On the other hand, if results turn out positive, one could conclude that the treatment was effective and that the tests were sensitive to the effect.

## PROCEDURE

The patients used in this study were colored, male schizophrenics admitted to the Veterans Administration Hospital, North Little Rock, Arkansas, between February, 1950, and October, 1950. Within a week after admission every colored patient who was suspected of schizophrenia was administered the battery of tests, which included, in addition to the mazes and mirror-tracing, the Rorschach and Draw-a-Person. Four copies of the Gardner Behavior Chart were independently filled out for each patient by a nurse and three psychiatric aides. Patients who were mentally defective or who were clearly not schizophrenic were eliminated.

The patients who were accepted for study were placed alternately in a control and experimental group. [2] The experimental patients attended a weekly group psychotherapy session conducted by the senior author; otherwise their treatment in the hospital did not differ systematically from that received by the control patients, who were treated in the routine manner. Three and a half to four months after his first psychological examination each patient was re-examined with the same battery. Several of the patients selected for study were

[1]Reviewed by the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the authors are a result of their own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

[2]The authors wish to express their appreciation to the management of the hospital for support and encouragement during this research project, and to Dr. A. D. Rosenthal, the ward physician whose cooperation was essential to carrying out the investigation.

lost during the course of the year as a result of unavoidable conflicts with hospital routine and early departure from the hospital. At present, 21 patients have completed the investigation cycle, 10 experimental and 11 control.

The psychotherapy group, into which the experimental patients were placed, was one which had been operating under the same leadership since June, 1949. The size of this group varied from 12 to 18 patients. Although at any one meeting a core of eight or nine old members was always present, every two or three weeks a new member would be present and an older one absent because of trial visit or discharge.

Psychodrama was the therapy technique most consistently used in this group. The atmosphere was permissive and tolerant, any form of spontaneity was encouraged, and whenever a patient showed an inclination to assume leadership, he was allowed to do so. Psychodramas were for the most part pointed at interpersonal relations involved in ward and hospital adjustment. Thus the immediate goal of therapy was to improve the social adjustment of the patients within the hospital. A large portion of each hour-and-a-half session was spent in open discussion; attitudes, expressions, ventilations were received passively by the therapist, although they were sometimes rejected by patients. With the exception of a brief summary after each meeting, in which the group activities of that session were translated into simple terms of interpersonal relations and social adjustment, the therapist was careful to avoid interpretations.

Several people other than patients were usually present at the group meetings. These included two hospital aides, sometimes nurses and physicians, psychological interns, and a female, colored volunteer worker, who sometimes functioned as an auxiliary ego in dramas. All of these people were required to become integral parts of the group whenever they were present. The junior author was also present in the group as an auxiliary ego. His contact with the patients, in addition to his individual examinations, was restricted to role-playing. [8]

[8]It is not likely that the positive results reported in this study are due to the increased familiarity and rapport with the examiner obtained during the group psychotherapy meetings. His role was strictly that of an observer who followed the instructions of the therapist in all psychodramas. It was the opinion of both the therapist and the assistant that very little relationship was being established between the assistant and patients during the sessions. Results with the Gardner Behavior Chart, to be reported in a later publication, show a reliable difference for the two groups of patients. This is a control of the familiarity factor, since the ward personnel who made out the ratings were equally familiar with all patients, and in most cases not aware of which were control and which experimental.

The Porteus Maze Test and the Mirror-Tracing Test have the usual characteristics of performance tests. The subject is faced with a definite, objective problem to be solved and needs a minimum of language to solve it. They differ from other tests of the same type, such as the Wechsler-Bellevue performance tests, form boards, sorting tests, etc., in that the record blank contains, in addition to abstractions from the subject's activity such as time, errors, and number of trials, a pattern of pencil marks which parallels closely the actual details of the movements used by the subject in solving the problem. This makes it possible to analyze at leisure the motor responses of the patient.

These tests also have the advantage of being ones with which it is fairly easy to estimate the degree of patient cooperation. And one of the strong points of the Porteus Mazes is that only a very small number of hospitalized, psychotic patients fail to cooperate.

The mazes were administered in the usual way [2]; qualitative, as well as quantitative, scores were calculated [3]. The quantitative score on the maze test is the usual score which is expressed in years of test age. The qualitative score is obtained by counting and weighting characteristics of the subject's pencil tracing, such as cutting corners, crossing lines, making a wavy line, and lifting the pencil from the paper after tracing of the maze is started.

The Mirror-Tracing Test was administered in the manner described in a previous publication [1], with the exception that, after the preliminary direct vision tracing of the six-pointed star, the subject was required to trace the pattern only twice, instead of five times. Performance on the mirror-tracing was considered a "failure" if the patient did not progress halfway around the star in 10 minutes

and if he persisted in attempts to quit after three urgings from the examiner. In both cases the test was stopped.

## RESULTS

Table 1 presents the mean test age on the Porteus Mazes for both groups at the first and second examination. For both groups, N is 10; one control patient's record was not used, because he refused to cooperate on the first maze test. He is the only patient who so far has refused to attempt the mazes. At his second testing he did cooperate, and made a T. A. of 8.5 years and a qualitative score of 7.80.

TABLE 1
MEAN PORTEUS MAZE TEST AGES

| Group | First Examination | | Second Examination | | $t$ ($df=9$) |
|---|---|---|---|---|---|
| | $M$ | $\sigma_m$ | $M$ | $\sigma_m$ | |
| Test | 10.45 | 0.92 | 12.95 | 1.14 | 3.750 |
| Control | 10.10 | 1.04 | 10.90 | 0.81 | 0.697 |
| $t$ ($df=18$) | 0.252 | | 1.465 | | |

The means in Table 1 show a substantial increase in score for the Test Group, after the interval of three and a half to four months of psychotherapy, whereas the scores for the Control Group show only a very slight increase. The $t$-ratios shown in the last column indicate that the difference for the Test Group is significant beyond the 1 per cent level of confidence. On the other hand, the $t$ of 0.697 for the Control Group indicates that the difference is not significant.

In the last row of Table 1 are the $t$-ratios calculated for the differences between means of the two groups at the first and at the second examination period. The small difference in mean test age at the first examination is insignificant; the $t$ of 1.465 suggests that the mean difference of 2.05 years in test age at the second testing may be meaningful, although it does not imply statistical significance to the 10 per cent level.

Table 2 presents the mean qualitative scores for Test and Control Groups. In previous publications the total qualitative score has been used as the measure of performance quality. In the present investigation the total score has been converted into an average score per maze blank by dividing each patient's total qualitative score by the number of maze blanks used. The means in Table 2 are the means of these ratios.

TABLE 2
MEAN QUALITATIVE MAZE SCORES

| Group | First Examination | | Second Examination | | $t$ ($df=9$) |
|---|---|---|---|---|---|
| | $M$ | $\sigma_m$ | $M$ | $\sigma_m$ | |
| Test | 4.36 | 0.92 | 2.15 | 0.29 | 2.508 |
| Control | 5.01 | 0.71 | 5.03 | 0.91 | 0.026 |
| $t$ ($df=18$) | 0.556 | | 3.021 | | |

The two means for the Test Group show a marked reduction in qualitative errors; the means for the Control Group are almost the same. The $t$-ratios in the right-hand column indicate a significance level between 5 and 2 per cent for the Test Group difference, while the Control Group difference is, of course, insignificant. The ratios in the last row show that the two groups were not significantly different at the first testing, but at the second testing were different beyond the 1 per cent level.

Table 3 presents results for each individual

TABLE 3
TOTAL TIME IN SECONDS AND FAILURES OF INDIVIDUAL PATIENTS ON MIRROR-TRACING

| Patient | First Exam. | Second Exam. |
|---|---|---|
| | *Test Group* | |
| Si | 585 | 230 |
| Wi | 570 | 234 |
| Sc | 635 | 226 |
| Jo | 395 | 223 |
| Le | Failure | 311 |
| Wr | Failure | 525 |
| Bu | Failure | Failure |
| He | Failure | 588 |
| Ri | Failure | Failure |
| Ro | Failure | 1171 |
| | *Control Group* | |
| Ne | 666 | 559 |
| De | 170 | 911 |
| Du | 345 | 206 |
| Ja | 1460 | Failure |
| Le | 619 | Failure |
| Hi | Failure | Failure |
| Lo | Failure | Failure |
| Ad | Failure | Failure |
| Ev | Failure | 1396 |
| Da | Failure | Failure |
| Ly | Failure | 1205 |

patient on the Mirror-Tracing Test. The total time scores in the table are the sums of the time in seconds for the two trials at tracing the star pattern by mirror vision. Since a large proportion of the patients failed the test, and since it is difficult to express a failure in sensible quantitative form, statistical analysis of the results has not been attempted. The significance of the raw data for individual patients is obvious to inspection.

Although six patients in each group failed at the first examination, at the second examination only two test patients failed, while six control patients failed. Any time score above 600 seconds has been found to be abnormally long for this test. Examination of the last column in Table 3 reveals that only three of the 10 patients in the Test Group failed or made abnormally long records. On the other hand, 9 of the 11 Control Group patients either failed or took an exceedingly long time to do the problem. The performance of none of the test patients declined from first to second testing; the performance of three of the control patients did decline on the second testing.

## DISCUSSION

The results presented above are in rather marked contrast to the disappointing and contradictory findings of recent studies in which attempts have been made to evaluate, by means of psychological tests, various forms of therapy with mental patients. In other investigations heavy emphasis has been placed on projective tests and on performance tests of the Wechsler-Bellevue type which require a relatively high degree of attention, concentration, and cooperation before the patient understands the nature of the problem which the examiner places before him. It is likely that a large part of the sensitiveness of the Porteus Mazes and Mirror-Tracing is a result of the fact that it is extremely easy for the patient to grasp what is expected of him, and consequently he makes an immediate, sincere attempt to do it.

While the problems presented by these two tests are easy to grasp, carrying out the solution is not so simple. A surprising number of apparently normal college students have difficulty solving both the mazes and mirror-tracing, and the quality of their performances is quite variable. This difficulty of execution is another factor which probably in part explains the sensitiveness of the mazes and mirror-tracing.

The mazes and mirror-tracing depend upon overt motor performance of the individual, and both permit a careful, detailed analysis of motor reactions. This phase of the total picture of the patient has been neglected in present-day clinical psychology. While the ideational productions of a patient are admittedly important in understanding and controlling him, and while an abstract intellectual level obtained on the Wechsler-Bellevue is highly significant and useful, it is obvious that social adjustment is in the end actually a matter of bodily movement. It is a well-recognized fact, for example, that many socially adjusted ex-patients still have the bizarre ideational content they had in the hospital. The important thing, however, is the extent to which this content spills out the efferent pathways into activity, talking, letter writing, violence, or any behavior which involves the patient in interpersonal conflict.

## SUMMARY

The Porteus Maze Test and the Mirror-Tracing Test were administered twice to two groups of hospitalized schizophrenics, with an interval of approximately four months between the two examinations. One group received group psychotherapy during the interval between testings; the other group of patients received routine treatment. Marked differences were found between the performances of the two groups on both tests at the second examination. The psychotherapy group improved in quantitative and qualitative scores on the mazes; performance on mirror-tracing also improved. The patients in the control group did not show significant evidence of improvement on either test.

It may be concluded from these results that the group psychotherapy used brings about changes in the patients in the direction of improved social adjustment, and that the maze test and mirror-tracing are sensitive to the changes.

The sensitivity of the Porteus Maze and the Mirror-Tracing Test to group psychother-

apy is explained as a consequence of three characterisitics of these tests: (1) the readiness with which patients understand the problem, (2) the difficulty of performing the movements necessary to solution, and (3) the record of performance obtained permits a detailed analysis of the motor reactions of the patient during problem solving.

*Received February 8, 1951.*

REFERENCES

1. PETERS, H. N. The mirror-tracing test as a measure of social maladaptation. *J. abnorm. soc. Psychol.*, 1946, **41**, 437–448.
2. PORTEUS, S. D. *The maze test and mental differences.* Vineland, N. J.: Smith, 1933.
3. PORTEUS, S. D. *Qualitative performance in the maze test.* Vineland, N. J.: Smith, 1942.
4. PORTEUS, S. D. Q scores, temperament, and delinquency. *J. soc. Psychol.*, 1945, **21**, 81–103.
5. PORTEUS, S. D., AND PETERS, H. N. Maze test validation and psychosurgery. *Genet. Psychol. Monogr.*, 1947, **36**, 3–86.

# PREDICTING CONTINUATION IN THERAPY BY MEANS OF THE RORSCHACH TEST[1]

LAWRENCE S. ROGERS AND JOANNE KNAUSS
VA MENTAL HYGIENE UNIT, DENVER, COLORADO
AND
KENNETH R. HAMMOND
UNIVERSITY OF COLORADO

ALL Mental Hygiene Clinics face the important question of the patient's motivation for treatment. The planning of treatment from the outset may depend heavily on this question, although desire, or motivation, for treatment is admittedly a rather ephemeral concept. The present study is based on the assumption that such motivation might well be indicated by how long the patient stayed in treatment. In the clinic at which the present study was carried out, approximately 35 per cent of the patients return for fewer than five interviews. A decision was made, therefore, to define as "not desiring treatment" those patients who returned for fewer than five interviews. Those patients who returned for five or more interviews were defined as "desiring treatment."

Since the Rorschach provides one possible tool for predicting motivation for treatment, the problem investigated directly was: Will the Rorschach test aid in predicting whether an outpatient will return for five or more interviews? The Rorschach data were examined in two ways: by clinical-impression judgments of the entire protocol, and by an analysis of 99 scores or "signs" obtained from the responses.

## METHOD AND RESULTS

One hundred and nine Rorschach records were taken from the files of the clinic, and all identifying material was removed from the protocols. The criterion groups did not differ from each other nor from the over-all clinic population as presented in the Annual Report of the clinic with respect to age, education, and diagnosis. The diagnostic classification of the patients is presented in Table 1.

TABLE 1

PERCENTAGE OF CASES IN EACH DIAGNOSTIC GROUP, BY CRITERION GROUPS AND ANNUAL REPORT

| Group | Five or More Interviews | Less Than Five Interviews | Total Sample (N=109) | Annual Report (N=436) |
|---|---|---|---|---|
| Neurotics | 59 | 50 | 57 | 61 |
| Character Disorder | 29 | 23 | 27 | 21 |
| Schizophrenics | 9 | 23 | 12 | 11 |
| Miscellaneous | 3 | 4 | 4 | 7 |

Three psychologists independently judged each Rorschach protocol in a "total impression" manner, as to whether or not the patient would return for five or more interviews. There were 87 cases who returned for five or more interviews, and 22 who returned for less than five interviews.[2] The psychologists making the judgments did not know how many cases were in each category but were warned that they should *not* assume that the categories would contain an equal number of cases.

[1]Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the authors are the result of their own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

[2]The reason for this rather lopsided split between the "five or more" interview group and the "less than five" group was that these same cases were to be used for a more important study on improved vs. unimproved cases. In order to make the split in the latter study more nearly at the median, it was necessary to use a less equivalent cut in this study. More equivalent cuts are reported later.

The psychologists independently wrote out their manifest criteria for judging whether the patients would stay in treatment. These criteria are presented in Table 2. Although the criteria were not compared until all the protocols had been judged, Table 2 indicates considerable agreement among the three judges. This was confirmed by the intercorrelations among the judges (A vs. B, +. 76; A vs. C, + .85; B vs. C, + .67). This interobserver agreement may be considered a measure of reliability.

TABLE 2

RORSCHACH CRITERIA USED BY PSYCHOLOGISTS FOR JUDGING WHETHER PATIENTS WOULD RETURN FOR MORE OR LESS THAN FIVE INTERVIEWS

| Less than Five Interviews | Five or More Interviews |
|---|---|
| 1. Rigidity | 1. *R* range 15–40 |
| 2. No *M* | 2. No rejections |
| 3. Perseveration | 3. *FC* and/or *CF*+, but no *C* |
| 4. Evasiveness | 4. Presence of *FC* or *FK* |
| 5. High *F*+ percentage and high number of populars | 5. Additionals |
| | 6. Presence of *M*+ |
| | 7. Ability to elaborate |
| | 8. Recovery on rejected cards or after shock |

Fourfold tables were set up and chi squares were calculated between the psychologists' judgments of plus five and minus five interviews and the criterion, namely, whether or not the patient actually did return for more than five interviews. None of these chi squares was significant, i.e., no predictions were better than chance. Results are presented in the first row of Table 3.

TABLE 3

PSYCHOLOGISTS' JUDGMENTS VS. THE CRITERION OF RETURN FOR FIVE OR MORE, OR LESS THAN FIVE, INTERVIEWS, WITH THE FACTOR OF IMPROVEMENT CONSIDERED

| | Judge A | Judge B | Judge C |
|---|---|---|---|
| Criterion vs. Judgments | $\chi^2$=0.110 P= .80 | $\chi^2$=0.910 P= .50 | $\chi^2$=0.056 P= .90 |
| Criterion (unimproved cases only) vs. Judgments | $\chi^2$=0.045 P= .90 | $\chi^2$=2.30 P= .20 | $\chi^2$=0.083 P= .80 |
| Criterion (improved cases +5, unimproved cases −5) vs. Judgments | $\chi^2$=0.125 P= .80 | $\chi^2$=0.279 P= .70 | $\chi^2$=0.101 P= .80 |

The above negative results led to the conjecture that the psychologists might have been influenced in their judgments by certain data from the Rorschach record which indicated that the patient would not be susceptible to psychotherapy and therefore was likely to terminate the therapy sessions early. Since no patient used in this series who had less than five interviews was judged "improved" by the therapists, it was certain that "improvement" was a variable correlated with "desire for treatment."

The role of "improvement" as a factor in the psychologists' judgment was evaluated first by computing chi squares only for those cases terminated as "unimproved." This "unimproved" group consisted of patients who had five or more interviews and patients who had four or fewer interviews.

A fourfold table was set up for this group ($N = 50$; five or more, 28; less than five, 22) and chi squares were calculated. None of these predictions was above chance. Results are presented in the second row of Table 3.

The role of improvement as a factor was also evaluated statistically by considering in the criterion group of more than five interviews, only those patients who had been terminated "improved" ($N = 59$), and by considering in the criterion group of less than five interviews, only those patients terminated "unimproved" ($N = 22$). That is, we tied the variable of "improvement in therapy" to the variable of "desire for treatment" in a one-to-one fashion in order to evaluate its role in the psychologists' judgments. However, when a fourfold table was set up, none of the chi squares was significant. These chi squares are presented in row 3 of Table 3. Thus, the psychologists were unable to predict from the Rorschach record which patients would continue in treatment. These predictions were unaffected by the subsequent improvement or lack of improvement in therapy.

Since the total-impression method was not successful, the problem next investigated was: Will the criterion groups show differences in

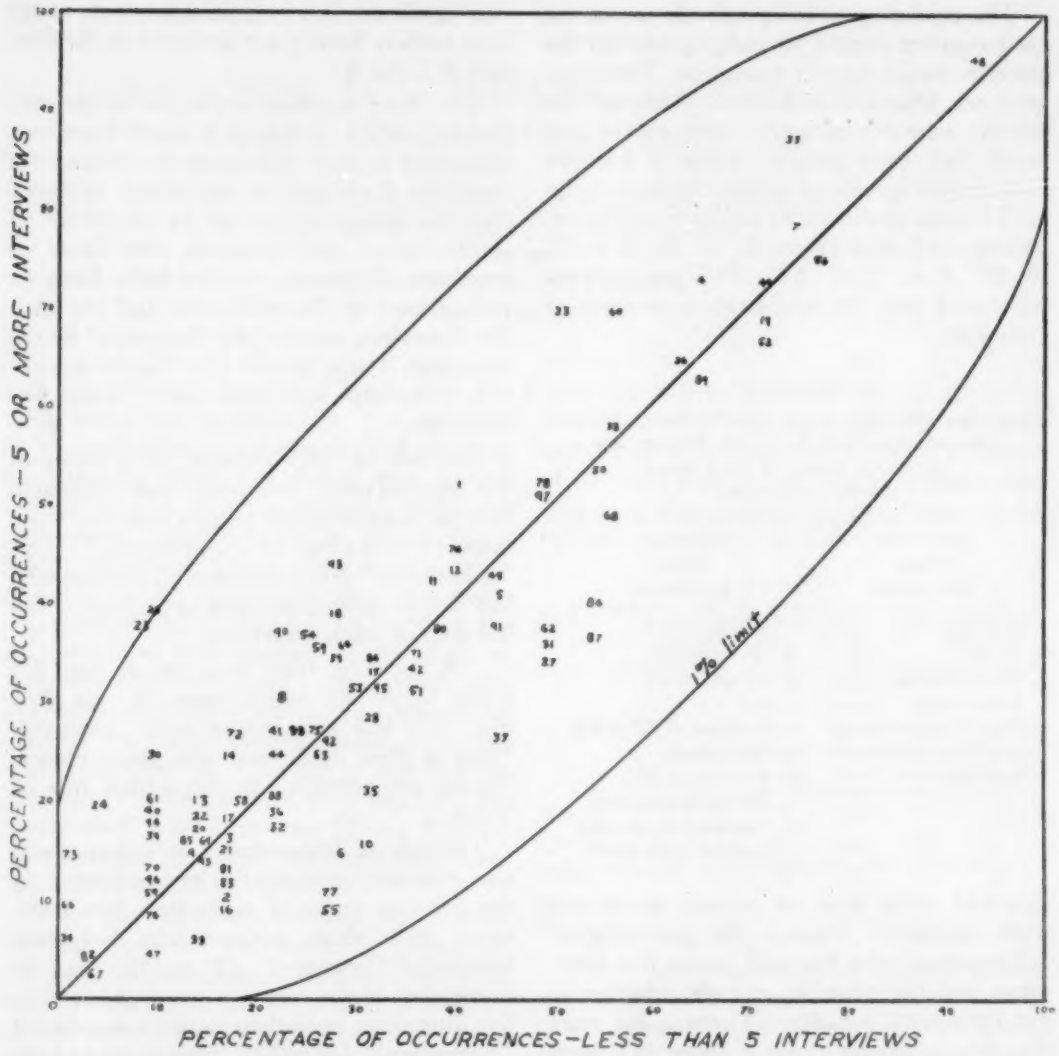


FIG. 1. Percentage of responses occurring in criterion groups. (See Table IV for responses indicated by number in this figure, e.g., $1 = R < 20$.)

the formal aspects of the Rorschach, e. g., differences in $M$, $\Sigma C$, etc. Significance tests were made on 99 formal scoring categories and combinations thereof. That is, the percentage of cases in each criterion group in which a given score or "sign" appeared was calculated, and the differences between these percentages were tested for significance. Results of these significance tests are presented pictorially in Figure 1. (Figure 1 is developed from a method presented by Zubin [1].) As Figure 1 indicates, differences between the criterion groups with respect to the formal scoring categories of the Rorschach are minimal, to say the least. (See Table 4 for the scoring categories indicated by the numbers in Figure 1.)

Whenever negative results are obtained with the Rorschach test, in which the skill of the psychologist plays such an important role, questions concerning the skill of the psychologists making the interpretations are bound to arise. The results of the 99 significance tests, however, confirm the results obtained by the total-impression method, for they indicate

TABLE 4
RORSCHACH "SIGNS" CORRESPONDING TO NUMBERS IN FIGURE 1

| | |
|---|---|
| 1. $R < 20$ | 51. Others presence |
| 2. $R < 10$ | 52. Categories $> 6$ |
| 3. $R \leqq 40$ | 53. Categories $> 8$ |
| 4. *M* presence | 54. Rejection presence |
| 5. $M > 1$ | 55. Rejections $\leqq 3$ |
| 6. $M > 3$ | 56. *At* presence |
| 7. *FM* presence | 57. $At > 2$ |
| 8. $FM \leqq 4$ | 58. $At > 3$ |
| 9. $FM > 4$ | 59. Sex presence |
| 10. *m* presence | 60. *N* or *Geog* presence |
| 11. Add. *m* presence | 61. *Art/Des* presence |
| 12. *k* presence | 62. Nature presence |
| 13. $k > 1$ | 63. *Geog* presence |
| 14. *K* presence | 64. Clouds presence |
| 15. *FK* presence | 65. Blood presence |
| 16. $FK \leqq 3$ | 66. Fire presence |
| 17. $F- \leqq 4$ | 67. $Ad > A$ |
| 18. $F- \leqq 3$ | 68. $F\% > 50$ |
| 19. *Fc* presence | 69. $F\% > 70$ |
| 20. $Fc \leqq 3$ | 70. $F\% < 30$ |
| 21. *c* presence | 71. $F\% \leqq 60$ |
| 22. *C'* presence | 72. $F\% < 40$ |
| 23. *FC* presence | 73. Shading absence |
| 24. $FC > 2$ | 74. Shading $\leqq 15\%$ |
| 25. $FC > 1$ | 75. Shading $\leqq 20\%$ |
| 26. *CF* presence | 76. Shading $\leqq 30\%$ |
| 27. $CF > 1$ | 77. $Sum\ C = 0$ |
| 28. $CF > 2$ | 78. $Sum\ C < 2$ |
| 29. $FC > CF$ | 79. $Sum\ C > 4$ |
| 30. *FC* presence & no *CF* | 80. $Sum\ C < 1.5$ |
| 31. $CF > FC$ | 81. $CF > FC$ & $Sum\ C > 3$ |
| 32. *CF* presence & no *FC* | 82. $M/Sum\ C = 0M/0C$ |
| 33. Presence of *CF* or *FC* | 83. $\%\ last\ 3 \geqq 24$ |
| 34. Presence of *C* | 84. $\%\ last\ 3 \geqq 30$ |
| 35. Human absence | 85. $\%\ last\ 3 \leqq 46$ |
| 36. Human $> 3$ | 86. $\%\ last\ 3 \leqq 40$ |
| 37. Human $> 2$ | 87. $\%\ C < 10$ |
| 38. *Hd* presence | 88. $\%\ C \leqq 20$ |
| 39. $Hd > 1$ | 89. $\%\ C \geqq 13$ |
| 40. $Hd > 3$ | 90. $W\% > 50$ |
| 41. $Hd > H$ | 91. $W\% \geqq 30$ |
| 42. $A\% \leqq 50$ | 92. $W\% < 20$ |
| 43. $A\% \leqq 60$ | 93. $W\% \leqq 45$ |
| 44. $A\% \leqq 30$ | 94. $D\% \geqq 30$ |
| 45. Populars $\leqq 5$ | 95. $D\% \geqq 40$ |
| 46. Populars $\leqq 2$ | 96. $D\% \leqq 65$ |
| 47. Populars $\leqq 8$ | 97. $D\% = 41$ to 60 |
| 48. $A > Ad$ | 98. *S* presence |
| 49. *A obj* presence | 99. $S > 1$ |
| 50. *Obj* presence | |

graphically how few differential cues were available to the psychologist in making his judgments.

## DISCUSSION

Our results indicate that in this situation, the Rorschach could not be used to predict whether or not a patient would stay in treatment for at least five interviews. Although it is recognized that it has never been claimed that the Rorschach test could be used to determine motivation for treatment, an opportunity was presented to investigate this problem. Desire for treatment, it should be pointed out, is a function not only of the patient's personality and his orientation to Mental Hygiene Clinic treatment, but also the skill with which the referral is made, as well as the type of treatment used at the particular clinic and the personality and skill of the therapist to whom he is assigned. It might be possible, therefore, within a given diagnostic category, with information concerning the type of referral and intake interview, and with a given therapist, to predict motivation for treatment successfully from the Rorschach test. It should be pointed out, however, that the investigation of a clinical testing procedure in a representative clinic situation is not without merit.

## SUMMARY

1. The Rorschach Records of 109 unselected Mental Hygiene Clinic patients were studied to determine whether it was possible to predict which of these patients would stay in treatment for five or more interviews.

2. Three psychologists, working independently, made clinical-impression judgments of the Rorschach protocol as to whether or not the patients would stay in treatment. Judgments were not better than chance.

3. The three psychologists showed considerable agreement in their selections and criteria.

4. The role of improvement as a factor in the experiment was evaluated and found to be irrelevant.

5. Statistical analysis of the formal scoring categories revealed only the number of significant differences expected by chance.

6. While these results are negative for an unselected clinic population, they do not preclude the possibility of securing positive results with more controls.

*Received January 16, 1951.*

## REFERENCES

1. ZUBIN, J. Note on a graphic method for determining the significant difference between group frequency. *J. educ. Psychol.*, 1936, **27**, 431–444.

# THE GENERALITY OF PSYCHIATRIC SYNDROMES[1]

J. R. WITTENBORN
YALE UNIVERSITY
AND
J. D. HOLZBERG
CONNECTICUT STATE HOSPITAL

THIS is one of a series of studies preliminary to the development of a quantified method for multiple psychiatric diagnosis [5]. The term "diagnosis" is here used in a limited descriptive sense, and the diagnostic method proposed is not concerned with etiological, prognostic, or dynamic considerations. It is proposed to devise a rating scale procedure whereby a psychologist, psychiatrist, nurse, or any other competent observer may rate the currently discernible symptoms of a mental hospital patient, score these ratings, and prepare a profile which would indicate the degree to which the patient's symptom manifestations resemble each of the various symptom patterns which may be shown to exist among mental hospital patients. The nature of the rating scales employed and the manner in which they were developed has been described elsewhere [7].

The feasibility of developing a quantified method for multiple psychiatric diagnosis is dependent upon the existence of clustering tendencies among psychiatric symptoms. It is recognized that a variety of circumstances can determine the existence as well as the nature of clustering tendencies. Nevertheless, it is believed that if clustering tendencies retain their essential characteristics for different raters and for patients in different hospitals, they could provide the basis for an effective multiple diagnostic procedure. In a preliminary study of 140 patients at the Northampton Veterans Hospital, the appearance of seven rather clear-cut symptom clusters was revealed by the method of multiple factor analysis [7]. The nature of the clustering was plausible from the standpoint of both the literature and clinical observations of patients' behavior.

It is the purpose of the present study either to confirm or to qualify the symptom clusterings revealed by the Northampton study. The Northampton data are not ideal for the purpose of revealing symptom clusters because no systematic sampling procedure determined the nature of the data included; the raters were asked to select for rating those patients whom they happened to know well. It is possible, therefore, that trivial or incidental characteristics of the Northampton sample could have produced the symptom clustering. The Northampton ratings were provided by psychiatric nurses as well as by psychiatrists of varying age, background, and point of view. The Northampton patients were quite varied; they included, for example, men from both World Wars and men whose illness was in some cases functional, in other cases related to alcoholism, and in other cases related to known brain pathology. Nevertheless, in such diversified data there was a conspicuous patterning of the symptoms.

In the present study the exact nature of the patient sample was carefully controlled. The psychiatrists at the Connecticut State Hospital, Middletown, Connecticut, rated all the patients admitted during the first six months of 1949. The ratings were a part of the psychiatrists' case preparation for the diagnostic staff conference. Each patient was presented at the diagnostic staff conference approximately three weeks after the date of his

[1]This research was made possible by the support of the U.S. Public Health Service; Mark A. May, Institute of Human Relations, Yale University; and Fred A. Mettler, Columbia University. The present data were provided by psychiatrists of the Connecticut State Hospital, including Benjamin Simon, former Clinical Director, E. K. Wilk, N. Toll, P. E. Yu, A. Santiccioli, C. E. Boyd, and other collaborators.

admission. The symptom ratings were based on observations of the patient made for the most part during the second week of his hospitalization. This was a very favorable arrangement because it meant that the ratings were made under relatively standardized conditions and, in so far as such things are feasible, at a standard time with respect to the management of the illness. None of the patients was under treatment during the rating period.

For the purposes of the present study, it was desired to restrict the analysis to patients who presented disturbances which were not believed to be due to toxic or neurological factors. Because of this, alcoholic, senile, sclerotic, or paretic patients were excluded.[2] Patients over sixty years of age were excluded, and psychopathic personalities were excluded because, as a hospitalized group, they are likely to be free of currently discernible pathology. The present restriction may be expected to affect the pattern of symptom clusters. As indicated in the discussion for the Northampton study [7], an analysis of symptom ratings for recently admitted, relatively young functional psychotics may emphasize the excited fulminating stage of mental illness and may not emphasize the affectively flattened, negativistic phases observed in patients who have been hospitalized for long periods.

If in the present sample of patients there is no conspicuous clustering of symptoms, or if there is a gross inexplicable difference between the symptom pattern found for the Northampton patients and for the Middletown patients, the value of basing a descriptive procedure on a patterning of symptoms is indeed questionable. If, however, the symptom pattern revealed in the Middletown study has essentially the same descriptive implications as the symptom pattern revealed in the Northampton study, we may conclude that there is a discernible tendency for symptoms to pattern and that this tendency is relatively independent of the differences in raters and the differences in patients which exist among many mental hospitals. Stated differently, if essentially the same symptom patterns are found in the two samples, it may be assumed that symptom-patterning is a sufficiently stable phenomenon to justify using it as a basis for a descriptive procedure.

[2]A small number of these patients were inadvertently included in the present sample.

## THE ANALYSIS

In order to study the symptom-patterning among the group of 250 Middletown patients, each of the symptom rating scales was correlated with every other one. Four of the 55 rating scales were rejected on statistical grounds (they were rarely rated at a pathological level, or they failed to yield significant correlations). The intercorrelations for the remaining 51 rating scales were submitted to a centroid analysis, and after the extraction of the seventh factor the residuals were small and showed no clustering tendencies. The centroid factors were submitted to a lengthy series of orthogonal rotations. The matrix of intercorrelations and the rotated factor matrix are given in Tables 1 and 2, respectively.[3] Although oblique reference vectors could pass more closely through the centroids of the various clusters than the present orthogonal reference vectors, it is apparent from examination of Table 2 that the orthogonal reference vectors are sufficient to indicate clearly the nature of the various symptom clusters and to indicate roughly the relative importance of the various symptoms in each of the clusters.

## DISCUSSION

The general implications of the present analysis are quite similar to the implications of the analysis performed on the Northampton data. There is a conspicuous tendency for the symptoms to cluster. The clusters are not only relatively clear-cut and characterized by relatively high factor loadings, but they are plausible both from the standpoint of the literature and from the standpoint of the impressions one gains from observations and clinical study of patients.

There are, however, certain differences between the cluster patterns revealed in the

[3]To reduce printing costs, the centroid factor matrix and the orthogonal transformation matrix have been deposited with the American Documentation Institute. Order Document 3371 from the American Documentation Institute, 1719 N Street, N.W., Washington 6, D.C., remitting $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film) or $1.00 for photocopies (6 × 8 inches) readable without optical aid.

TABLE 1
CORRELATION MATRIX

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Acute insomnia | | | | | | | | |
| 2. Ideas change with spontaneous rapidity | 357 | | | | | | | |
| 3. Unjustified sexual beliefs | 000 | 139 | | | | | | |
| 4. Cannot banish obsessive thoughts | 020 | 011 | 146 | | | | | |
| 5. Delusional belief that he is evil | 157 | —056 | 000 | 294 | | | | |
| 6. Gives in easily to others | —135 | —289 | —032 | 202 | 246 | | | |
| 7. In almost constant movement | 386 | 636 | 071 | 037 | —024 | —298 | | |
| 8. Unaware of the feelings of others | 311 | 346 | —058 | —034 | 014 | —120 | 319 | |
| 9. Use made of physical disease symptoms | 121 | —018 | 125 | —016 | —042 | 108 | 048 | —154 |
| 10. Refuses to eat | 406 | 158 | —059 | —023 | 119 | —015 | 113 | 301 |
| 11. Deliberately disrupts routines | 352 | 409 | 071 | —027 | —071 | —239 | 485 | 376 |
| 12. Temper tantrums | 324 | 359 | 108 | 052 | —027 | —309 | 452 | 224 |
| 13. Avoids people | 129 | —092 | 000 | 135 | 297 | 253 | —225 | 256 |
| 14. Shouts, sings, and talks loudly | 345 | 554 | 089 | 057 | —105 | —324 | 668 | 274 |
| 15. Behavior disrupted by phobias | 096 | 000 | 066 | 399 | 133 | 017 | 164 | 076 |
| 16. Incontinent because of own negligence | 444 | 336 | —031 | —073 | 111 | —098 | 348 | 509 |
| 17. Engrossed in plans | 167 | 408 | 344 | —022 | 058 | —119 | 350 | —026 |
| 18. Feelings of impending doom | 185 | —026 | 119 | 254 | 523 | 077 | 034 | —041 |
| 19. Cannot resist compulsive acts | (Omitted) | | | | | | | |
| 20. Exaggeration of ability and well-being | 068 | 436 | 164 | 021 | —137 | —198 | 395 | 126 |
| 21. Unable to stick to or carry out any plan | 201 | 323 | 030 | 059 | 046 | 079 | 277 | 504 |
| 22. Cannot believe that he can be helped | 200 | —088 | 000 | 253 | 542 | 192 | —054 | 013 |
| 23. Fears others misunderstand him | 015 | —032 | 108 | 182 | 273 | 188 | 041 | —217 |
| 24. Patient's thinking clearly delusional | 134 | 212 | 248 | 022 | 057 | —189 | 131 | 238 |
| 25. No organic basis for complaints | 150 | 043 | 144 | 056 | —036 | 138 | 083 | 011 |
| 26. Feels systematically persecuted | 009 | 010 | 270 | 070 | —011 | —129 | —026 | —010 |
| 27. Believes others influence him | 000 | 105 | 343 | 256 | 093 | 055 | 000 | 049 |
| 28. Desperately distressed by his anxiety | 161 | —049 | 055 | 248 | 294 | 202 | 074 | —154 |
| 29. Organic pathology with emotional basis | 103 | —037 | 028 | 075 | —070 | 129 | —042 | —064 |
| 30. Rigidly orderly | —126 | —058 | 044 | 118 | 088 | 204 | —083 | —261 |
| 31. Dramatically attention-demanding | 140 | 260 | 185 | 024 | —093 | —192 | 305 | 170 |
| 32. All overt activity is at a minimum | 093 | —075 | —084 | 131 | 241 | 366 | —182 | 299 |
| 33. Grandiose notions | 016 | 285 | 151 | —015 | —057 | —118 | 144 | 122 |
| 34. Does not believe he has a problem | 080 | 230 | 097 | —138 | —236 | —286 | 228 | 358 |
| 35. Compulsive acts continuous | 115 | 144 | 139 | 397 | 187 | 024 | 195 | 086 |
| 36. Great variation occurs in rate of speech | 227 | 538 | 039 | 015 | 039 | —142 | 491 | 361 |
| 37. Initiates physical assaults | 424 | 386 | 031 | —048 | —047 | —242 | 461 | 401 |
| 38. Delusions of homosexual attacks | (Omitted) | | | | | | | |
| 39. Mood changes very frequent and abrupt | 335 | 459 | 032 | 150 | 130 | —134 | 529 | 149 |
| 40. Has made attempts at suicide | 137 | —162 | —033 | 193 | 450 | 207 | —063 | —107 |
| 41. Failures of affective response | —059 | 102 | 143 | 011 | —058 | 105 | —076 | 331 |
| 42. No concern over physical handicaps | 033 | —053 | 000 | —093 | 141 | 083 | 046 | 037 |
| 43. Cannot make decisions | 232 | 110 | —041 | 080 | 156 | 258 | 087 | 430 |
| 44. Opinions exceptional to physical laws | 217 | 258 | 116 | —056 | 015 | —318 | 233 | 215 |
| 45. Hallucinations | 108 | 141 | 093 | 058 | 094 | —029 | 000 | 276 |
| 46. Memory faults | 087 | 141 | —106 | —102 | —080 | 096 | 090 | 377 |
| 47. Fear of committing an abhorred act | 133 | —047 | 213 | 207 | 188 | 139 | —041 | —197 |
| 48. Words not relevant to recognizable idea | 099 | 279 | 090 | —023 | —075 | 015 | 137 | 483 |
| 49. Shows failure and blocking | 123 | 102 | 033 | 153 | 265 | 039 | 025 | —030 |
| 50. Repudiates earlier insights | 108 | 295 | 142 | —009 | —012 | —074 | 145 | 131 |
| 51. Speech is stilted | 052 | 370 | 042 | 162 | —042 | —108 | 207 | 192 |
| 52. Overt homosexual demands | (Omitted) | | | | | | | |
| 53. Lies or steals | (Omitted) | | | | | | | |
| 54. Exaggerated affective expressions | 371 | 460 | 056 | 011 | 056 | —189 | 448 | 090 |
| 55. Characteristically oppositional | 451 | 403 | 078 | —040 | 013 | —216 | 435 | 503 |

TABLE 1 (Continued)

| | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | | | | | | | | | | | | | | | |
| 2. | | | | | | | | | | | | | | | |
| 3. | | | | | | | | | | | | | | | |
| 4. | | | | | | | | | | | | | | | |
| 5. | | | | | | | | | | | | | | | |
| 6. | | | | | | | | | | | | | | | |
| 7. | | | | | | | | | | | | | | | |
| 8. | | | | | | | | | | | | | | | |
| 9. | | | | | | | | | | | | | | | |
| 10. | —059 | | | | | | | | | | | | | | |
| 11. | 023 | 117 | | | | | | | | | | | | | |
| 12. | 144 | 051 | 518 | | | | | | | | | | | | |
| 13. | —046 | 313 | —078 | —119 | | | | | | | | | | | |
| 14. | 059 | 070 | 503 | 525 | —214 | | | | | | | | | | |
| 15. | 000 | —063 | 019 | 071 | 110 | 189 | | | | | | | | | |
| 16. | —084 | 404 | 267 | 136 | 175 | 284 | 067 | | | | | | | | |
| 17. | 153 | —082 | 259 | 297 | —223 | 299 | 061 | 000 | | | | | | | |
| 18. | 099 | 099 | —017 | 153 | 263 | 000 | 189 | 075 | 027 | | | | | | |
| 19. | (Omitted) | | | | | | | | | | | | | | |
| 20. | 036 | —078 | 432 | 400 | —212 | 492 | —029 | 000 | 501 | —026 | | | | | |
| 21. | —020 | 245 | 206 | 065 | 202 | 187 | 145 | 397 | 139 | 043 | 132 | | | | |
| 22. | 038 | 094 | —064 | —012 | 267 | —121 | 180 | 057 | —026 | 484 | —198 | 110 | | | |
| 23. | 120 | —120 | —041 | 078 | 060 | —017 | 229 | —164 | 265 | 137 | 000 | —017 | 131 | | |
| 24. | 051 | 118 | 119 | 238 | 254 | 127 | 061 | 183 | 210 | 163 | 100 | 214 | 009 | —011 | |
| 25. | 481 | —034 | —027 | 062 | 000 | 171 | 102 | 024 | 110 | 000 | —042 | 000 | 087 | 055 | 051 |
| 26. | 090 | —043 | 155 | 282 | 063 | 053 | —036 | —091 | 145 | 171 | 118 | —098 | 061 | 091 | 399 |
| 27. | 088 | —012 | 149 | 170 | 161 | 063 | 014 | —066 | 072 | 200 | 115 | 038 | 036 | 167 | 337 |
| 28. | 092 | —039 | 000 | 120 | 046 | 106 | 397 | —126 | 025 | 343 | —097 | —027 | 277 | 305 | —051 |
| 29. | 386 | 052 | —031 | 036 | —046 | 066 | 000 | 056 | —025 | 052 | —024 | —054 | 063 | —016 | —084 |
| 30. | 174 | —145 | —049 | —075 | —096 | —041 | —023 | —263 | 199 | 062 | 038 | —148 | 039 | 377 | —133 |
| 31. | 308 | —058 | 313 | 451 | —136 | 394 | 081 | 015 | 339 | 044 | 404 | 134 | 056 | 071 | 150 |
| 32. | —075 | 187 | —111 | —194 | 437 | —133 | 060 | 227 | —232 | 147 | —221 | 273 | 280 | —081 | 069 |
| 33. | 100 | —018 | 234 | 146 | 021 | 125 | —120 | 057 | 208 | 018 | 428 | 128 | —171 | 000 | 345 |
| 34. | —129 | 146 | 192 | 209 | 018 | 230 | —068 | 326 | 074 | —184 | 127 | 196 | —227 | —261 | 306 |
| 35. | 000 | 000 | 052 | 060 | 129 | 242 | 393 | 094 | 043 | 088 | 000 | 113 | 147 | 188 | —014 |
| 36. | 000 | 185 | 397 | 253 | 065 | 426 | 083 | 374 | 143 | 056 | 204 | 303 | —035 | —023 | 143 |
| 37. | 021 | 338 | 596 | 535 | 051 | 383 | 000 | 344 | 114 | 088 | 189 | 286 | —056 | —036 | 217 |
| 38. | (Omitted) | | | | | | | | | | | | | | |
| 39. | 021 | 107 | 255 | 445 | —071 | 490 | 188 | 179 | 267 | 199 | 253 | 157 | 117 | 093 | 148 |
| 40. | —066 | 063 | 000 | —043 | 129 | —141 | 052 | 000 | —061 | 377 | —116 | —016 | 405 | 057 | —071 |
| 41. | —076 | 190 | —081 | —186 | 493 | —136 | 000 | 231 | —105 | —054 | —175 | 278 | —078 | —066 | 324 |
| 42. | 320 | 038 | 068 | 069 | —022 | 057 | —106 | 161 | 073 | —076 | 070 | 078 | —091 | —023 | —073 |
| 43. | 000 | 304 | 047 | —107 | 344 | —019 | 121 | 375 | —152 | 049 | —126 | 552 | 187 | —084 | 195 |
| 44. | 058 | —014 | 444 | 414 | 032 | 248 | —031 | 073 | 267 | 138 | 304 | 070 | 053 | 000 | 381 |
| 45. | —037 | 115 | 148 | 081 | 363 | 027 | 119 | 113 | 034 | 151 | 033 | 199 | 051 | 043 | 439 |
| 46. | —053 | 225 | 060 | —023 | 190 | —063 | —070 | 451 | —169 | —075 | —092 | 332 | —048 | —167 | 112 |
| 47. | 178 | —025 | —060 | 000 | 177 | —025 | 169 | —054 | 049 | 228 | 046 | 026 | 193 | 154 | 016 |
| 48. | —119 | 142 | 118 | —039 | 249 | 000 | —016 | 482 | —110 | —057 | —052 | 362 | —149 | —190 | 310 |
| 49. | 109 | 031 | 111 | 042 | 055 | 015 | 174 | 016 | 181 | 312 | 057 | 127 | 238 | 303 | 090 |
| 50. | 063 | —023 | 228 | 297 | 000 | 180 | 038 | 048 | 174 | 146 | 269 | 150 | —064 | —027 | 268 |
| 51. | 055 | —059 | 165 | 234 | —023 | 356 | 044 | 021 | 229 | —020 | 364 | 202 | —019 | 144 | 127 |
| 52. | (Omitted) | | | | | | | | | | | | | | |
| 53. | (Omitted) | | | | | | | | | | | | | | |
| 54. | 112 | 080 | 364 | 410 | —109 | 597 | 163 | 239 | 308 | 146 | 390 | 054 | 101 | 048 | 128 |
| 55. | —017 | 406 | 499 | 369 | 158 | 369 | 069 | 549 | 147 | 061 | 091 | 402 | 070 | —075 | 237 |

TABLE 1 (Continued)

| | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 39 | 40 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | | | | | | | | | | | | | | | |
| 2. | | | | | | | | | | | | | | | |
| 3. | | | | | | | | | | | | | | | |
| 4. | | | | | | | | | | | | | | | |
| 5. | | | | | | | | | | | | | | | |
| 6. | | | | | | | | | | | | | | | |
| 7. | | | | | | | | | | | | | | | |
| 8. | | | | | | | | | | | | | | | |
| 9. | | | | | | | | | | | | | | | |
| 10. | | | | | | | | | | | | | | | |
| 11. | | | | | | | | | | | | | | | |
| 12. | | | | | | | | | | | | | | | |
| 13. | | | | | | | | | | | | | | | |
| 14. | | | | | | | | | | | | | | | |
| 15. | | | | | | | | | | | | | | | |
| 16. | | | | | | | | | | | | | | | |
| 17. | | | | | | | | | | | | | | | |
| 18. | | | | | | | | | | | | | | | |
| 19. | | | | | | | | | | | | | | | |
| 20. | | | | | | | | | | | | | | | |
| 21. | | | | | | | | | | | | | | | |
| 22. | | | | | | | | | | | | | | | |
| 23. | | | | | | | | | | | | | | | |
| 24. | | | | | | | | | | | | | | | |
| 25. | | | | | | | | | | | | | | | |
| 26. | —035 | | | | | | | | | | | | | | |
| 27. | 000 | 513 | | | | | | | | | | | | | |
| 28. | 171 | 080 | 223 | | | | | | | | | | | | |
| 29. | 552 | —010 | 000 | 110 | | | | | | | | | | | |
| 30. | 134 | 110 | 110 | 193 | 096 | | | | | | | | | | |
| 31. | 225 | 122 | 143 | 027 | 163 | —021 | | | | | | | | | |
| 32. | 043 | —111 | 024 | 025 | 000 | —117 | —152 | | | | | | | | |
| 33. | —029 | 163 | 334 | —134 | —050 | 000 | 204 | —068 | | | | | | | |
| 34. | —006 | 117 | —013 | —302 | —043 | —203 | 088 | 029 | 136 | | | | | | |
| 35. | 071 | —017 | 078 | 144 | —102 | 065 | 069 | 063 | —028 | —012 | | | | | |
| 36. | 030 | 000 | —016 | —052 | —034 | —082 | 173 | 070 | 118 | 202 | 232 | | | | |
| 37. | —060 | 168 | 131 | —014 | —055 | —194 | 214 | 000 | 187 | 201 | 023 | 407 | | | |
| 38. | (Omitted) | | | | | | | | | | | | | | |
| 39. | 125 | 116 | 095 | 143 | 100 | —022 | 254 | —072 | 000 | 142 | 216 | 484 | 320 | | |
| 40. | —127 | 012 | —014 | 250 | 000 | 000 | —114 | 209 | —060 | —206 | 049 | 062 | 016 | 034 | |
| 41. | 011 | 062 | 181 | —153 | —127 | —200 | —113 | 335 | 034 | 200 | 000 | 143 | 028 | —148 | —152 |
| 42. | 262 | 029 | 017 | —089 | 407 | 056 | 138 | 018 | 072 | 072 | —119 | 000 | 079 | 021 | —127 |
| 43. | 048 | —112 | —035 | 046 | —018 | —130 | —081 | 549 | 025 | 118 | 092 | 220 | 195 | 021 | 044 |
| 44. | —045 | 356 | 196 | —129 | —064 | —081 | 401 | —144 | 333 | 218 | 000 | 181 | 360 | 210 | 031 |
| 45. | 007 | 249 | 402 | 108 | —082 | —052 | 018 | 084 | 282 | 077 | 028 | 082 | 213 | 010 | —079 |
| 46. | —020 | —152 | —011 | —211 | 047 | —147 | —026 | 320 | 000 | 320 | —019 | 280 | 170 | 054 | —028 |
| 47. | 082 | 038 | 135 | 308 | 094 | 111 | 026 | 120 | 000 | —138 | 079 | 000 | 000 | —027 | 254 |
| 48. | —023 | 011 | 075 | —173 | 013 | —229 | 015 | 323 | 054 | 340 | 022 | 355 | 163 | 108 | —095 |
| 49. | 025 | 012 | 180 | 365 | 044 | 137 | 048 | 044 | 119 | —093 | 122 | 103 | 081 | 102 | 139 |
| 50. | 073 | 257 | 210 | 032 | 147 | —016 | 187 | 000 | 257 | 117 | —106 | 104 | 200 | 221 | —038 |
| 51. | 080 | 135 | 141 | —037 | —037 | 000 | 369 | 019 | 201 | 097 | 186 | 208 | 124 | 129 | —155 |
| 52. | (Omitted) | | | | | | | | | | | | | | |
| 53. | (Omitted) | | | | | | | | | | | | | | |
| 54. | 075 | 030 | 000 | 025 | 123 | —039 | 399 | —126 | 152 | 137 | 145 | 332 | 277 | 520 | —015 |
| 55. | 040 | 065 | 120 | —069 | 034 | —216 | 242 | 187 | 156 | 349 | 077 | 405 | 682 | 308 | —055 |

TABLE 1 (Continued)

| | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 | 54 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | | | | | | | | | | | | |
| 2. | | | | | | | | | | | | |
| 3. | | | | | | | | | | | | |
| 4. | | | | | | | | | | | | |
| 5. | | | | | | | | | | | | |
| 6. | | | | | | | | | | | | |
| 7. | | | | | | | | | | | | |
| 8. | | | | | | | | | | | | |
| 9. | | | | | | | | | | | | |
| 10. | | | | | | | | | | | | |
| 11. | | | | | | | | | | | | |
| 12. | | | | | | | | | | | | |
| 13. | | | | | | | | | | | | |
| 14. | | | | | | | | | | | | |
| 15. | | | | | | | | | | | | |
| 16. | | | | | | | | | | | | |
| 17. | | | | | | | | | | | | |
| 18. | | | | | | | | | | | | |
| 19. | | | | | | | | | | | | |
| 20. | | | | | | | | | | | | |
| 21. | | | | | | | | | | | | |
| 22. | | | | | | | | | | | | |
| 23. | | | | | | | | | | | | |
| 24. | | | | | | | | | | | | |
| 25. | | | | | | | | | | | | |
| 26. | | | | | | | | | | | | |
| 27. | | | | | | | | | | | | |
| 28. | | | | | | | | | | | | |
| 29. | | | | | | | | | | | | |
| 30. | | | | | | | | | | | | |
| 31. | | | | | | | | | | | | |
| 32. | | | | | | | | | | | | |
| 33. | | | | | | | | | | | | |
| 34. | | | | | | | | | | | | |
| 35. | | | | | | | | | | | | |
| 36. | | | | | | | | | | | | |
| 37. | | | | | | | | | | | | |
| 38. | | | | | | | | | | | | |
| 39. | | | | | | | | | | | | |
| 40. | | | | | | | | | | | | |
| 41. | | | | | | | | | | | | |
| 42. | —092 | | | | | | | | | | | |
| 43. | 351 | 000 | | | | | | | | | | |
| 44. | 027 | 000 | —067 | | | | | | | | | |
| 45. | 309 | —084 | 223 | 200 | | | | | | | | |
| 46. | 255 | 185 | 314 | 061 | 079 | | | | | | | |
| 47. | 024 | —068 | 053 | —074 | 080 | —112 | | | | | | |
| 48. | 493 | —019 | 406 | 125 | 196 | 453 | —025 | | | | | |
| 49. | 060 | —063 | 141 | 137 | 167 | —028 | 140 | —031 | | | | |
| 50. | 087 | 015 | 094 | 287 | 227 | 080 | 041 | 272 | 212 | | | |
| 51. | 095 | 027 | —069 | 232 | 050 | —018 | 036 | 040 | 066 | 079 | | |
| 52. | (Omitted) | | | | | | | | | | | |
| 53. | (Omitted) | | | | | | | | | | | |
| 54. | —219 | 036 | —037 | 299 | 033 | —035 | 024 | —067 | 088 | 191 | 242 | |
| 55. | 226 | 083 | 335 | 289 | 287 | 383 | —022 | 322 | 123 | 207 | 173 | 371 |

TABLE 2
ROTATED FACTOR MATRIX

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | $H^2$ |
|---|---|---|---|---|---|---|---|---|
| 1. Acute insomnia | - 96 | *526* | 223 | *341* | 224 | - 91 | - 19 | 511 |
| 2. Ideas change with spontaneous rapidity | - 4 | *475* | *517* | -112 | - 20 | *303* | 153 | 621 |
| 3. Unjustified sexual beliefs | *330* | - 65 | 160 | 55 | 84 | 174 | 144 | 200 |
| 4. Cannot banish obsessive thoughts | 144 | -123 | 8 | *336* | - 56 | 51 | *487* | 392 |
| 5. Delusional belief that he is evil | - 47 | 45 | -126 | *722* | 18 | 154 | 31 | 566 |
| 6. Gives in easily to others | -143 | -159 | *-427* | 178 | 183 | 219 | 142 | 361 |
| 7. In almost constant movement | -125 | *441* | *626* | - 15 | 10 | 33 | 253 | 667 |
| 8. Unaware of the feelings of others | 116 | *696* | - 48 | -163 | - 22 | 82 | 173 | 564 |
| 9. Use made of physical disease symptoms | 30 | -135 | 153 | 82 | *581* | 130 | - 78 | 410 |
| 10. Refuses to eat | - 72 | *520* | -133 | 144 | 105 | -121 | - 87 | 347 |
| 11. Deliberately disrupts routines | 167 | *435* | *479* | 0 | 66 | - 58 | 18 | 455 |
| 12. Temper tantrums | 255 | 271 | *604* | 77 | 131 | -109 | 54 | 541 |
| 13. Avoids people | 245 | *306* | *-496* | 298 | 26 | 121 | 71 | 509 |
| 14. Shouts, sings, and talks loudly | - 18 | *334* | *692* | - 64 | 79 | 6 | 254 | 666 |
| 15. Behavior disrupted by phobias | 0 | - 1 | 32 | 245 | 34 | - 43 | *587* | 409 |
| 16. Incontinent because of own negligence | -121 | *739* | - 34 | 36 | 102 | - 15 | 50 | 576 |
| 17. Engrossed in plans | 59 | - 26 | *511* | 54 | 128 | *404* | 3 | 448 |
| 18. Feelings of impending doom | 186 | 57 | 4 | *646* | 63 | 42 | 59 | 464 |
| 19. Cannot resist compulsive acts | (Omitted) | | | | | | | |
| 20. Exaggeration of ability and well-being | 154 | 63 | *610* | - 91 | - 68 | *358* | - 64 | 545 |
| 21. Unable to stick to or carry out any plan | - 14 | *528* | - 89 | - 3 | 27 | *351* | 249 | 473 |
| 22. Cannot believe that he can be helped | - 26 | 47 | -129 | *654* | 30 | 34 | 97 | 459 |
| 23. Fears others misunderstand him | - 26 | -270 | 100 | *301* | 156 | 202 | 219 | 287 |
| 24. Patient's thinking clearly delusional | *554* | *318* | 31 | 20 | 19 | 227 | - 11 | 461 |
| 25. No organic basis for complaints | - 34 | - 69 | 68 | - 56 | *658* | 61 | 189 | 486 |
| 26. Feels systematically persecuted | *639* | - 56 | 157 | 57 | 88 | - 28 | - 48 | 450 |
| 27. Believes others influence him | *645* | - 26 | 18 | 109 | 121 | 66 | 189 | 484 |
| 28. Desperately distressed by his anxiety | - 21 | -174 | 29 | *441* | 207 | - 83 | *372* | 414 |
| 29. Organic pathology with emotional basis | - 69 | - 55 | 28 | - 6 | *676* | - 36 | - 4 | 467 |
| 30. Rigidly orderly | - 61 | -383 | 67 | 96 | 240 | 150 | 98 | 254 |
| 31. Dramatically attention-demanding | 211 | 97 | *490* | 7 | 199 | 195 | 38 | 373 |
| 32. All overt activity is at a minimum | - 33 | *337* | *-498* | 164 | 1 | 214 | 151 | 458 |
| 33. Grandiose notions | *405* | 109 | 202 | - 52 | 34 | *316* | -178 | 352 |
| 34. Does not believe he has a problem | 186 | *421* | 101 | -357 | - 50 | - 1 | 5 | 352 |
| 35. Compulsive acts continuous | - 18 | 40 | 121 | 194 | - 96 | 22 | *519* | 333 |
| 36. Great variation occurs in rate of speech | - 57 | *514* | *317* | - 7 | - 30 | 95 | 265 | 448 |
| 37. Initiates physical assaults | 175 | *620* | *317* | 77 | 87 | -217 | - 9 | 576 |
| 38. Delusions of homosexual attacks | (Omitted) | | | | | | | |
| 39. Mood changes very frequent and abrupt | - 38 | 299 | *508* | 140 | 84 | 8 | 261 | 444 |
| 40. Has made attempts at suicide | - 41 | 7 | - 88 | *564* | -140 | - 56 | - 26 | 351 |
| 41. Failures of affective response | 241 | *329* | *-434* | -185 | - 22 | 286 | 140 | 491 |
| 42. No concern over physical handicaps | - 7 | 56 | 41 | - 93 | *476* | - 2 | -155 | 264 |
| 43. Cannot make decisions | - 85 | *531* | *-398* | 57 | 85 | 247 | 210 | 563 |
| 44. Opinions exceptional to physical laws | *429* | 255 | *378* | 84 | - 14 | 120 | -168 | 442 |
| 45. Hallucinations | *487* | 249 | -150 | 83 | 78 | 76 | 98 | 350 |
| 46. Memory faults | - 49 | *497* | -255 | -212 | 128 | 58 | 40 | 381 |
| 47. Fear of committing an abhorred act | 89 | - 89 | - 47 | *347* | 130 | 115 | 132 | 186 |
| 48. Words not relevant to recognizable idea | 131 | *533* | -240 | -279 | - 7 | 248 | 170 | 527 |
| 49. Shows failure and blocking | 52 | 28 | 13 | *349* | 179 | 221 | 173 | 236 |
| 50. Repudiates earlier insights | *307* | 186 | 169 | - 40 | 145 | 216 | 16 | 227 |
| 51. Speech is stilted | 118 | 102 | *315* | - 33 | - 15 | *316* | 154 | 248 |
| 52. Overt homosexual demands | (Omitted) | | | | | | | |
| 53. Lies or steals | (Omitted) | | | | | | | |
| 54. Exaggerated affective expressions | - 38 | 274 | *596* | 182 | 101 | 79 | 18 | 482 |
| 55. Characteristically oppositional | 115 | *767* | 170 | 40 | 174 | - 88 | 90 | 678 |

Northampton data and those revealed in the present data. The differences appear to be related to known differences between the two samples and in this sense serve to strengthen the hypothesis that the patterning of symptoms is an orderly, dependable phenomenon and not capricious or haphazard. For example, the present sample provides no evidence for the symptom-clustering which in the study of Northampton patients was considered to be due to the presence of patients who had been long hospitalized and presented symptoms reminiscent of the hebephrenic or deteriorated designations. Inasmuch as the Middletown sample comprises relatively young patients recently admitted to the hospital, the types of behavior implied by the terms hebephrenic and deteriorated are not likely to be encountered sufficiently often to determine a separate pattern of intercorrelations.

It is also noted from Table 2 that the excited cluster is much more conspicuous in the Middletown sample than it is in the Northampton sample. This cluster is particularly conspicuous and involves a relatively large number of the rating scales. This is considered to be related to the fact that the Middletown sample included relatively young functional patients who for the most part were in the initial fulminating phases of their illness.

There is another respect in which the two factor analyses are different from each other, and this difference also is taken as an expression of a known difference between the samples. In the Middletown sample one factor, factor 7, is determined primarily by obsessive thinking, phobias, anxiety, and compulsive behavior. Although no great significance is ascribed to this factor, a pattern of phobic-compulsive reaction is likely to be more conspicuous among newly admitted patients than among patients who have been hospitalized for some period of time. Such behavior is not absent from patients who have been hospitalized for long periods of time, but it is not one of their most conspicuous characteristics.

In the present analysis the loadings for factors 1, 2, 3, 4, 5, and 6 are evidence for the Paranoid Schizophrenic, Excited, Manic-Depressed, Anxiety, Hysterical, and Paranoid Condition symptom clusters, respectively. The implication of these clusters in the present sample are considered to be essentially the same as those for the respectively designated clusters of the Northampton sample. These two samples despite differences in raters, differences in location, and differences in sources of patients, present highly similar clusters of symptoms; this fact is taken as an evidence that the clustering of symptoms is a fairly stable phenomenon and may justifiably be employed as a basis for descriptive procedure.

Although symptom-clustering may be a relatively stable phenomenon, the forces which determine it remain unknown. It is quite possible that the clustering is due to patterns of belief and technique of examination prevalent among contemporary New England psychiatrists. Possibly, psychiatrists who are different in cultural background or training would rate patients so differently as to result in the appearance of different symptom patterns. It is also possible that the forms of behavior by which mental illness is manifested are in some way culturally determined; this possibility can be examined only by making intercultural comparisons for symptom patterns among the mentally ill.

## SUMMARY

The present investigation comprises a factor analysis of the intercorrelations among 51 symptom rating scales. The ratings were based on a sample of 250 patients recently admitted to the Connecticut State Hospital at Middletown. In so far as possible, the sample was restricted to patients under 60 years of age suffering from a functional disorder.

The purpose of the present investigation is to examine the possibility that the clustering of symptoms among mental hospital patients is relatively stable and not greatly altered by differences in raters and differences in hospitals. The present analysis revealed eight relatively clear-cut symptom clusters; in the order of the factors presented in Table 2 they are tentatively described by the following terms: Paranoid Schizophrenia, Excitement, Manic state–Depressed state, Anxiety, Hysteria, Paranoid Condition, and Phobic-Compulsive reaction. The implications of these clusters are considered to be in excellent agreement with the implications of clusters revealed in a factor analysis based on ratings for pa-

tients at the Northampton Veterans Hospital. Differences between the two analyses are accounted for in terms of major differences between the two samples. The consistency between these two analyses is taken as evidence that symptom-clustering is a sufficiently stable phenomenon to justify its use as a basis for descriptive procedures.

*Received November 13, 1950.*

## References

1. Kraepelin, E. *General paresis.* New York: J. nerv. ment. Dis. Publishing Co., 1913.
2. Kraepelin, E. *Dementia praecox.* Edinburgh: E. & S. Livingstone, 1919.
3. Kraepelin, E. *Manic-depressive insanity and paranoia.* Edinburgh: E. & S. Livingstone, 1921.
4. Moore, T. V. *The essential psychoses and their fundamental syndromes.* Baltimore: Williams & Wilkins, 1933.
5. Wittenborn, J. R. A new procedure for evaluating mental hospital patients. *J. consult. Psychol.,* 1950, **14**, 500–501.
6. Wittenborn, J. R., and Mettler, F. A. Some psychological changes following psychosurgery. *J. abnorm. soc. Psychol.* (In press.)
7. Wittenborn, J. R. Symptom patterns in a group of mental hospital patients. *J. consult. Psychol.,* 1951, **15**, 290-302.

# EQUIVALENT TEST AND MENTAL AGES FOR THE WISC

DAVID WECHSLER
NEW YORK UNIVERSITY

ONE of the important features of the Wechsler Intelligence Scale for Children has been the abandonment of mental age both as a method for calculating IQ's and for defining levels of intelligence. There were several reasons [1] why this was done, the main one being the hope that it would discourage the practice of comparing individuals in terms of their mental ages, with the MA interpreted as an absolute measure of level of intelligence. This interpretation, in the author's opinion, is not valid. However, since the MA has such a firm place in clinical practice and is often a legal necessity, several methods whereby equivalent "MA's" (actually, test ages) may be derived from the WISC are presented here for those who wish to use them, in spite of the fact that these methods violate the philosophy of the Scale.

There are various ways in which an "MA" may be obtained from the weighted scores furnished by the WISC. Only three will be discussed. The first and most direct one may be termed the *Formula* method. This is obtained by treating the MA as the unknown in the formula IQ = MA/CA. Since the WISC furnishes an IQ and the child's chronological age is known, the MA is computed by multiplying the subject's obtained IQ by his CA. Thus, if a child's IQ is 115 and his chronological age 85 months, his MA would be 98 months or 8 years 2 months.

The second and third procedures are, respectively, the *Mean Test-Age* and *Median Test-Age* methods. For these, a table of equivalent test-ages for all raw scores on each of the subtests is necessary. Table 1 gives these equivalents.[1] It has been derived as follows: For each age interval, the raw score corresponding to the scaled score of 10 has been taken to represent mean test performance for that particular age, as in point of fact it is. Table 1 records test-ages for each subtest and for each age interval. The table should be used as follows: Locate an individual's subtest raw score in the column which is headed with the name of the subtest. The corresponding test-age is the entry, in the same row, in the column headed "Test Age" at either end of the table. The raw score represents the performance level of the average child at that age.

Example: Subject S (irrespective of age) makes the following scores: Information 9, Comprehension 6, Arithmetic 8, and so on. Looking up the table, we find that an Information score of 9 is made by the average child in the age group 7 years 10 months (approximate mid-point of the interval 7 years 8 months to 7 years 11 months), a Comprehension score of 6 by a child 6 years 6 months (approximate mid-point of the interval 6 years 4 months to 6 years 7 months), and so on. The raw scores for all the other subtests are treated in the same manner.

In looking up equivalent test-ages for specific test scores, occasional lacunae will be encountered: (1) The same raw score will be found opposite several successive test ages. If the raw score applies to only two successive ages, use the lower age. If there are more than two, use a median value. (2) The obtained raw scores are beyond the limits found in the table. In such instances one arbitrarily assigns the age equivalent *4 years 10 months* (or,

[1]Test-age equivalents between ages 8 years 2 months through 12 years 2 months for raw scores on Code A and test-age equivalents from 5 years 10 months to 7 years 10 months for raw scores on Code B are extrapolated values.

TABLE 1

TEST AGE EQUIVALENTS FOR WISC RAW SCORES

| Test Age | Information | Comprehension | Arithmetic | Similarities | Vocabulary | Digit Span | Picture Completion | Picture Arrangement | Block Design | Object Assembly | Code A | Code B | Mazes | Test Age |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5-2 | 5 | 5 | 3 | 3 | 15 | 5 | 6 | — | 4 | 8–9 | 15–16 | — | 5 | 5-2 |
| 5-6 | — | 5 | — | 3 | 15 | 5 | 6 | 4 | 4 | 8–9 | 17–20 | — | 5 | 5-6 |
| 5-10 | 6 | — | 4 | 4 | 16–17 | 6 | 6 | 5 | — | 10–11 | 21–24 | 10 & below | 6 | 5-10 |
| 6-2 | — | — | 4 | 4 | 16–17 | 6 | — | 6 | — | 10–11 | 25–27 | 11–12 | 7 | 6-2 |
| 6-6 | 7 | 6 | — | — | 18 | 7 | 7 | 7–9 | 5 | 12–13 | 28–31 | 13–16 | 8 | 6-6 |
| 6-10 | — | 7 | 5 | 5 | 19–20 | — | 7 | 10–11 | — | 12–13 | 32–33 | 17–20 | 9 | 6-10 |
| 7-2 | 8 | 7 | 5 | 5 | 19–20 | — | — | 12–13 | 6 | 14–15 | 33–34 | — | 10 | 7-2 |
| 7-6 | — | 8 | — | 6 | 21–22 | 8 | 8 | 14–16 | 6 | 14–15 | 35–37 | 21–22 | 11 | 7-6 |
| 7-10 | 9 | — | 6 | 6 | 23–24 | 8 | — | 17–19 | 7–8 | 15–16 | 35–37 | 23–24 | 12 | 7-10 |
| 8-2 | — | 9 | — | — | 23–24 | — | — | 20–21 | 8–9 | 16–17 | 38–39 | 25–26 | 13 | 8-2 |
| 8-6 | 10 | 9 | 7 | 7 | 25–26 | — | 9 | 22–23 | 9–10 | 17–18 | 40–41 | 25–26 | 14 | 8-6 |
| 8-10 | 11 | 10 | — | 7 | 26–27 | — | 9 | 23–24 | 10–11 | 17–18 | — | 27–29 | 14 | 8-10 |
| 9-2 | 11 | 10 | — | 8 | 27–28 | 9 | 10 | 24–25 | 11–12 | 19–20 | 42–43 | 27–29 | 15 | 9-2 |
| 9-6 | — | 11 | 8 | 8 | 28–29 | 9 | 10 | 25–26 | 13–15 | 19–20 | 44–45 | 30–32 | 15 | 9-6 |
| 9-10 | 12 | 11 | 8 | 8 | 29–30 | 9 | 10 | 25–26 | 13–15 | 19–20 | — | 33–35 | 15 | 9-10 |
| 10-2 | 12 | 12 | — | 9 | 31–32 | 9 | 11 | 27–28 | 16–20 | 21–22 | 46 | 33–35 | 16 | 10-2 |
| 10-6 | 13 | 12 | 9 | 9 | 32–33 | — | 11 | 27–28 | 16–20 | 21–22 | — | 36–37 | 16 | 10-6 |
| 10-10 | 13 | 12 | 9 | 9 | 34–35 | — | 11 | 27–28 | 16–20 | 21–22 | — | 38–39 | 16 | 10-10 |
| 11-2 | 14 | 13 | 10 | 10 | 34–35 | 10 | 12 | 29 | 21–24 | 23 | 47 | 40 | 17 | 11-2 |
| 11-6 | 14 | 13 | 10 | 10 | 36–37 | 10 | 12 | 29 | 21–24 | 23 | — | 41 | 17 | 11-6 |
| 11-10 | 15 | 14 | 10 | 10 | 38 | 10 | 12 | 29 | 21–24 | 23 | — | 42–43 | 17 | 11-10 |
| 12-2 | 16 | 15 | 11 | 11–12 | 39 | 10 | — | 30 | 25–28 | 24 | 48 | 44 | 17 | 12-2 |
| 12-6 | 17 | 15 | 11 | 11–12 | 39 | 10 | — | 30 | 29–30 | 24 | — | 45–46 | 17 | 12-6 |
| 12-10 | 17 | 15 | 11 | 11–12 | 40–41 | 10 | — | 30 | 29–30 | 24 | — | 47 | 17 | 12-10 |
| 13-2 | 18 | 16 | 12 | 13 | 42 | 10 | 13 | 31–33 | 31 | 24 | — | 48–49 | 17 | 13-2 |
| 13-6 | 18 | 16 | 12 | 13 | 43–44 | 10 | 13 | 31–33 | 32–33 | 24 | — | 48–49 | 17 | 13-6 |
| 13-10 | 18 | 16 | 12 | 13 | 43–44 | 10 | 13 | 31–33 | 32–33 | 24 | — | 50–51 | 17 | 13-10 |
| 14-2 | 19 | 16 | 12 | 13 | 45–46 | 11 | 13 | 34 | 34–35 | 25 | — | 52–53 | 17 | 14-2 |
| 14-6 | 19 | 16 | 12 | 13 | 45–46 | 11 | 13 | 34 | 34–35 | 25 | — | 54–55 | — | 14-6 |
| 14-10 | 19 | 17 | 12 | 14 | 45–46 | 11 | 13 | 34 | 34–35 | 25 | — | 54–55 | — | 14-10 |
| 15-2 | 20 | 17 | — | 14 | 47–49 | 11 | 14 | 35–36 | 36–39 | 26 | — | 56–57 | 18 | 15-2 |
| 15-6 | 20 | 17 | — | 14 | 47–49 | 11 | 14 | 35–36 | 36–39 | 26 | — | 56–57 | 18 | 15-6 |
| 15-10 | 20 | 17 | — | 14 | 47–49 | 11 | 14 | 35–36 | 36–39 | 26 | — | 56–57 | 18 | 15-10 |

merely, below 5 years) for scores falling below the lower limit and *15 years 10 months* for scores beyond the upper limit.[3]

[3]In preparing the data for the correlations cited below, the method was slightly altered; see text.

To obtain a mean test (or "equivalent mental") age, sum the several test-ages and divide by the number of tests given. To obtain a median test-age, distribute the test-age scores and find the central value in the usual way. The difference in values obtained by the two

methods will usually be small. If ten tests are used the Mean method is generally the easier. An example of the calculations involved in all methods is given in Table 2.

TABLE 2

COMPUTATION OF "MA" EQUIVALENTS FROM WISC SCORES

| Male, Age 10 years 6 months, WISC IQ 89 | | |
|---|---|---|
| Test | Raw Score | Test Age |
| Information | 11 | 8–10 |
| Comprehension | 10 | 8–10 |
| Arithmetic | 9 | 10–6 |
| Similarities | 7 | 8–6 |
| Vocabulary | 34 | 10–10 |
| (Digit Span) | 8 | 7–6 |
| Picture Completion | 11 | 10–6 |
| Picture Arrangement | 20 | 8–2 |
| Block Design | 14 | 9–6 |
| Object Assembly | 19 | 9–6 |
| Coding | 21 | 7–6 |
| (Mazes) | 16 | 10–6 |

1. *Formula Method*

   $MA = IQ \times CA = .89 \times 126 \text{ mos.} = 112.14 \text{ mos.} = 9 \text{ yrs. } 4 \text{ mos.}$

2. *Mean Method*

   $\frac{\Sigma(\text{Test-age Sc.})}{N} = \frac{104 \text{ yrs. } 80 \text{ mos.}}{12} =$ 9 yrs. 2 mos.

3. *Median Method*

   $\frac{8 \text{ yrs. } 10 \text{ mos.} + 9 \text{ yrs. } 6 \text{ mos.}}{2} = 9 \text{ yrs. } 2 \text{ mos.}$

A comparison of the correlations between the various methods is given in Table 3. These are reported separately for 100 boys at age 10-6 and 100 girls at 13-6, both taken from the standardizing population. The test-age range in both these samples was from 5 years 2 months to 15 years 10 months. Some of the obtained raw scores fell outside this range; the procedure followed in such instances was to assign a test-age of *4 years 10 months* to all cases below the lower end of the table, and *16 years 2 months* to all cases above.

Perusal of Table 3 shows high intercorrelation between the three methods used for calculating mental ages from WISC scores. The highest average correlation is furnished by the Formula method, which also tends to give somewhat higher mental age equivalents. The latter is noticeably true for the older age

TABLE 3

INTERCORRELATIONS OF THREE METHODS OF DETERMINING MENTAL AGES FROM WICS SCORES

| | Mean* Method | Median† Method | Formula‡ Method |
|---|---|---|---|
| *A. 100 boys, average age 10–6, from the standardization population* | | | |
| Median Method | .920 | | |
| Formula Method | .969 | .913 | |
| Mean (in months) | 128.4 | 127.5 | 128.9 |
| SD | 21.4 | 21.7 | 19.8 |
| *B. 100 girls, average age 13–6, from the standardization population* | | | |
| Median Method | .968 | | |
| Formula Method | .971 | .949 | |
| Mean (in months) | 153.0 | 155.9 | 159.6 |
| SD | 23.5 | 26.7 | 25.6 |

*MA = mean of test-age equivalents
†MA = median of test-age equivalents
‡MA = IQ × CA (in months)

group, as might be expected.[3] The Median Test-Age method produces somewhat lower intercorrelations with the Formula method than does the Mean method. This may be due, in part, to the fact that there are only 10 subtests; a larger number of subtests might have stabilized the median and given higher correlations. Another slight advantage of the Mean Test-Age method is that it follows closely the procedure used in derivation of IQ's in standardization.[4] The CA × IQ formula correlates very highly with both Median and Mean methods.

## SUMMARY AND DISCUSSION

Three methods have been furnished for obtaining mental age equivalents from the WISC. Two of the methods require a table of equivalent test-age scores. This table also permits an examiner to obtain age norms for a given test performance. The author does not favor using MA equivalents for intelligence test scores, but he does see a legitimate use for test-age equivalents if they are employed to show how a child of a given age compares with children of his own age in performance on a given test. In such cases, however, one should remember that the test age obtained is to be

[3]This is due to the test-age ceiling or highest subtest equivalents possible (15 years 10 months) on any of the subtests.

[4]The derivation of IQ's is based on the total of standard scores for the subtests, a statistic which correlates perfectly with mean standard score.

interpreted not as a measure of intelligence but as a measure of a specific aptitude.

*Received June 12, 1951.*
*Early Publication.*

## Reference

1. Wechsler, D. *The Wechsler Intelligence Scale for Children.* New York: Psychological Corp., 1949.

# SOME EVIDENCE ON THE VALIDITY OF THE WISC[1]

JOHN J. PASTOVIC AND GEORGE M. GUTHRIE
THE PENNSYLVANIA STATE COLLEGE

BEFORE the Wechsler Intelligence Scale for Children (WISC) [6] can be substituted for accepted tests, its correlation with them should be determined at appropriate age levels. Four unpublished studies on this topic and the results of the present study are presented in Table 1. It is apparent that the Binet IQ exceeds the WISC by greater amounts at lower age levels. This increasing discrepancy casts doubt on the sampling procedure used to obtain the WISC standardization groups.

This study investigates the relationship of the WISC and the Revised Stanford-Binet, Form L, with preschool and second-grade children. Fifty children were tested at each of two age levels, five years six months, and seven years six months. The order of administration was alternated to control practice effect.

Our results tend to contradict Seashore's contention based on an analysis of the standardization data that, "*WISC* will approximate in meaning (as far as size of the number is concerned) the IQ's secured by the Stanford-Binet Revision" [5, p. 110]. In addition, our data and those of three other studies cited [1, 2, 4], find a higher mean performance IQ than verbal IQ over a wide range of intelligence, although the WISC was intended to have a difference of zero points between these averages.

We conclude that the WISC IQ should not be interpreted as equivalent to a Binet IQ at age levels below ten years since the WISC score is consistently lower than that of the Binet.

*Received May 14, 1951.*
*Early Publication.*

[1]This article is based on a Master's thesis of the first writer under the direction of the second writer.

TABLE 1

SUMMARY OF VALIDITY STUDIES OF THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

| Investigator | Test† | Mean IQ | *CR*‡ | Correlation with Criterion |
|---|---|---|---|---|
| Clarke [1] | *Binet* | 96.12 | — | — |
| $N = 85$ | WISC V | 96.11 | 0.01 | .83 |
| $CA = 11\text{–}1$ | WISC P | 99.75 | 2.40* | .57 |
| 5th grade | WISC FS | 97.64 | 1.46 | .79 |
| McBrearty [2] | *Arthur* | 101.79 | — | — |
| $N = 52$ | WISC V | 95.83 | 2.72** | .55 |
| $CA = 11\text{–}2$ | WISC P | 99.00 | 1.41 | .65 |
| 5th grade | WISC FS | 97.12 | 2.57* | .71 |
| Nale [3] | *Binet* | 55.38 | — | — |
| $N = 104$ | | | | |
| $CA = 14\text{–}1$ | | | | |
| Defectives | WISC FS | 57.97 | 6.60** | .91 |
| Rapaport [4] | *Binet* | 97.01 | — | — |
| $N = 100$ | WISC V | 89.94 | 6.43** | .79 |
| $CA = 7\text{–}6$ | WISC P | 91.60 | 4.29** | .74 |
| Public school | WISC FS | 89.60 | 7.80** | .85 |
| Pastovic | *Binet* | 115.08 | — | — |
| $N = 50$ | WISC V | 108.56 | 6.39** | .82 |
| $CA = 7\text{–}6$ | WISC P | 112.68 | 1.92 | .71 |
| 2nd grade | WISC FS | 111.50 | 4.41** | .88 |
| Pastovic | *Binet* | 113.02 | — | — |
| $N = 50$ | WISC V | 101.58 | 7.78** | .63 |
| $CA = 5\text{–}6$ | WISC P | 104.24 | 4.85** | .57 |
| Kindergarten | WISC FS | 103.16 | 7.41** | .71 |

†Criterion test is in italics.
‡The critical ratio between the criterion and the WISC scales.
*P < .05.
**P < .01.

REFERENCES

1. CLARKE, F. R. A comparative study of the WISC and Revised Stanford-Binet Intelligence Scale, Form L, in relation to scholastic achievement of a fifth grade population. Unpublished

Master's thesis, The Pennsylvania State College, 1950.

2. McBREARTY, J. F. A comparison of the WISC with the Arthur Performance Scale, Form I, and their relationship to the Progressive Achievement Tests. Unpublished Master's thesis, The Pennsylvania State College, 1951.
3. NALE, STANLEY L. The Wechsler Intelligence Scale for Children and the Revised Stanford-Binet on 104 defective patients at the Polk State School. Paper read at the Amer. Ass. on Ment. Def., New York, May, 1951.
4. RAPAPORT, I. A comparison of performance on the Wechsler Intelligence Scale for Children and the Revised Stanford-Binet Scale. Unpublished Master's thesis, University of Pittsburgh, 1951.
5. SEASHORE, H., WESMAN, A., AND DOPPELT, J. The standardization of the Wechsler Intelligence Scale for Children. *J. consult. Psychol.*, 1950, **14**, 99–110.
6. WECHSLER, D. *Wechsler Intelligence Scale for Children.* New York: Psychological Corp., 1949.

# THE REGRESSION OF SCHIZOPHRENIC LANGUAGE[1]

ROBERT B. ELLSWORTH
THE PENNSYLVANIA STATE COLLEGE

THIS paper is an evaluation of intellectual and emotional regression in schizophrenic patients. The schizophrenic's habits of personal cleanliness, his private and magical language, and his lack of integration with consensual reality, often seem childlike when compared with normal adult behavior. Because of this similarity, many psychologists and psychiatrists have interpreted the overt behavior of schizophrenics as an emotional and intellectual return to childhood.

The concept of schizophrenic regression has gained theoretical importance in the field of psychoanalysis. Landis and Bolles [11, p. 51] state: "The psychoanalyst believes that this regression is a real psychological return to an emotional and intellectual infantilism." Fenichel [9, p. 415] describes schizophrenic regression as reaching back much earlier than does any regression in neurosis "to the time when the ego first came into being."

On the other hand, Cameron [6, pp. 491-492] delineates the concept of regression by stating: "In the sense of our definition of regression, as a reversion to ways or objects which during earlier phases of one's development were accepted as adequate and appropriate, it is obvious that schizophrenic techniques frequently deserve to be called regressive. However, the older designation of schizophrenia as a 'regressive psychosis' which implied sometimes a return to childhood. . . is not justified by clinical facts."

Previous studies have indicated that language differences are highly significant for the interpretation of both intellectual and emotional adjustment. The intellectual development of the child is characterized by changes in the percentages of nouns, verbs, adjectives, and pronouns in his spoken and written vocabulary [1, 12]. Balken and Masserman [2], Busemann [4], and Boder [3] have found that the adjective-verb quotient is an indicator of emotional stability. That language differences can be used to differentiate schizophrenics from normal adults has been shown by Fairbanks [8].

## THE PROBLEM

If schizophrenics do regress toward childhood, the language they utilize as part of their overt behavior should reflect this regression. The percentages of nouns, verbs, pronouns, and adjectives used by the schizophrenic group in their language will first be compared with those used by a normal adult control group. If there are significant differences in the parts-of-speech percentages between these two groups, the parts-of-speech percentages of the schizophrenic group will then be compared with the percentages used by children's groups at various age levels. This procedure will indicate whether or not there is a regression in the language of schizophrenics and, if so, just how far this regression extends.

## METHOD OF PROCEDURE

*Material used to measure the language differences.* Language variability between large groups has been measured adequately with written samples of language [12], and because of inadequate physical facilities and the large number of subjects, it was decided that written material rather than actual verbatim transcriptions should be employed.

Of three methods tried in a preliminary study with a group of third- and fifth-grade children, it was found that the sentence completion test was the only method sensitive

[1]The present study is part of a Master's thesis [7] completed at the Utah State College. The data were compiled at the Idaho State Mental Hospital.

enough to differentiate significantly all parts-of-speech percentages between the two groups.[2]

The criteria of a noun, pronoun, verb, and adjective were taken from the definitions of these parts of speech in *Webster's Collegiate Dictionary* [15].

*Subjects.* Each of the groups—normal, schizophrenic, and control—contained from 30 to 50 subjects. The *normal* groups were chosen from public school students in the third, fifth, seventh, ninth, and eleventh grades. In addition, language samples were obtained from college students between the ages of 18 and 27. The *mental patient* group was selected on the basis of two criteria: (1) a psychiatric diagnosis of schizophrenia and (2) obvious signs of overinclusion and disorganization in their language. The mean age and environmental background of the schizophrenic group and control group were matched as nearly as possible. The *control* group was selected from much the same environment as that in which the schizophrenic patients had lived, i.e., a farming district. The average age of the control group was 41.3 years, and the average age of the mental patient group was 42.6 years.

TABLE 1

PERCENTAGES OF THE PARTS OF SPEECH FOR THE EXPERIMENTAL GROUPS ON THE BASIS OF THE SENTENCE COMPLETION TEST

| Group | Percentages of the Parts of Speech | | | |
|---|---|---|---|---|
| | Nouns | Pronouns | Verbs | Adjectives |
| 5th Grade | 27.6 | 19.2 | 30.5 | 21.9 |
| 7th Grade | 26.8 | 18.1 | 30.0 | 25.1 |
| 9th Grade | 26.9 | 18.2 | 30.1 | 26.2 |
| 11th Grade | 26.5 | 18.9 | 32.2 | 24.2 |
| College | 25.7 | 18.6 | 31.6 | 25.6 |
| Control | 25.5 | 16.3 | 31.6 | 27.0 |

*Administration of the test.* A set of instructions was read to each group of subjects before they were given the test. The subjects of school age were tested in the classroom, whereas the schizophrenics and the control subjects were tested in smaller groups.

[2]The other two methods used in obtaining written samples of language were essay-type questions. On one, the subject was asked to write a story beginning, "Once upon a time." On the other type, the subject was asked to finish a story about a doctor, the introduction of which was structured for him.

The third-grade children were dropped from this study. During the test administration they asked so many questions about procedure and spelling, often rewriting their sentences with different words rather than waiting to have their original word choice spelled for them, that it was thought unlikely their sentence completion tests offered a valid indication of their use of the parts of speech. The widely scattered distribution of parts-of-speech percentages for this particular group supports this conjecture.

## RESULTS

Table 1 indicates the change in parts-of-speech percentages with age. It is evident that children's language contains more nouns and pronouns and fewer verbs and adjectives than the language of older normal subjects.

Table 2 presents the parts-of-speech percentages of the fifth-grade, schizophrenic, and control groups. Between the fifth-grade group and the control group, the noun and adjective percentage differences were significant at the 1 per cent level of confidence and the pronoun percentage difference was significant at the 5 per cent level, but the verb percentage difference between the two groups was not statistically significant.

TABLE 2

PERCENTAGES OF THE PARTS OF SPEECH FOR THE FIFTH-GRADE, SCHIZOPHRENIC, AND NORMAL CONTROL GROUPS

| Group | Percentages of the Parts of Speech | | | |
|---|---|---|---|---|
| | Nouns | Pronouns | Verbs | Adjectives |
| 5th Grade | 27.6 | 19.2 | 30.5 | 21.9 |
| Schizophrenic | 27.5 | 20.1 | 30.9 | 22.0 |
| Control | 25.5 | 16.3 | 31.6 | 27.0 |

No significant parts-of-speech percentage differences were found between the fifth-grade children and the schizophrenic group. No other group showed as much similarity in its parts-of-speech percentages as the fifth-grade group. All other groups compared with the schizophrenic group had at least two of the four parts of speech differing significantly at the 5 per cent level of confidence.

## DISCUSSION

The percentage of nouns seems to reflect

the extent of concretistic thinking. Piaget [13], in particular, has pointed out that the child's thinking is concretistic as evidenced by the high percentage of nouns.

In a recent article Goldstein [10] interprets schizophrenic thinking as concretistic in nature, and he points out that this type of thinking hinders one's ability to plan ahead ideationally and to assume an attitude toward the "mere" possible. Because of concretistic thinking, both the child and the schizophrenic should have difficulty in planning and in refining their language and thinking in terms of a possible listener. As Sullivan [14, p. 12] states: "The refinement [of language] is largely the reference to other people as potential hearers."

The child and the schizophrenic assume that other people know what they are talking about because they themselves know what they mean; hence, they see no need to clarify and refine their language. Sullivan [14] illustrated this type of thinking with a story about a boy and his "possible" pony. The child wanted a pony and was promised one by his parents. However, because the child's concept of the pony was private and unshared, he expected a pony with qualities nonexistent and assumed that his parents also knew of these qualities. This type of concept formation is readily observable in the schizophrenic patient.

Adjectives in the language of both the schizophrenic and children seem to be related to their inability to refine and communicate their thoughts. Since adjectives describe and clarify, they are not used extensively by the child or the schizophrenic because neither sees any need to clarify his language.

The relative lack of verbs, although statistically insignificant, signifies a trend toward emotional instability [3]. Both the schizophrenic and the child are probably less stable emotionally than the normal adult.

Both the child and the schizophrenic used more pronouns than the normal subject. Several authors have pointed out that this is related to the egocentricity of the child and the schizophrenic because it indicates self-reference or "I-ness."

It was found in this particular study that the schizophrenic group used fewer personal pronouns than did either the control group or the fifth-grade group. Instead, the schizophrenic group used many third-person pronouns, and in so doing were able to project many ideas and possible problems onto other objects and persons. This would seem to indicate that the theory of schizophrenic egocentricity based on previous studies of their language might possibly have been based on hasty assumptions and needs further investigation. The schizophrenic is probably more concerned with projecting painful material away from himself than he is in drawing attention to himself.

*Concept of Regression.* About two-thirds of the schizophrenic group indicated disorganization in their writing. The other third organized their written language into socially meaningful form so that it could not be discriminated from the organization in the normal control group's written language. Therefore, the group was separated for further study. It must be remembered, however, that all the original schizophrenic group indicated schizophrenic thinking by their language when they were selected for this study. Table 3 gives the parts-of-speech percentages for the normal, fifth-grade, disorganized schizophrenic, and the organized schizophrenic groups.

TABLE 3

PERCENTAGES OF THE PARTS OF SPEECH FOR THE CONTROL GROUP, "ORGANIZED SCHIZOPHRENIC," FIFTH-GRADE, AND "DISORGANIZED SCHIZOPHRENIC" GROUPS

| Group | Percentages of the Parts of Speech | | | |
|---|---|---|---|---|
| | Nouns | Pronouns | Verbs | Adjectives |
| Control | 25.5 | 16.3 | 31.6 | 27.0 |
| Organized Schizophrenic | 23.0 | 18.5 | 33.0 | 25.0 |
| 5th Grade | 27.6 | 19.2 | 30.5 | 21.9 |
| Disorganized Schizophrenic | 29.3 | 20.6 | 29.7 | 21.0 |

The disorganized schizophrenics "regressed" from the organized schizophrenics, as indicated by the parts-of-speech percentages, in exactly the same manner as the fifth-grade group regressed from the normal control group. The organized schizophrenic group used approximately the same parts-of-speech percentages in their sentence completions as the normal control group. The disorganized schizophrenic

group "regressed" even further in their language than did the fifth-grade children by using more nouns and pronouns, and fewer verbs and adjectives than the fifth-grade group.

It might be assumed from these findings that language regression is a function of the inability to organize thought processes into socially acceptable and understandable form and that disorganization occurs when the individual withdraws socially and neglects to clarify his language in terms of a potential listener.

The same thing is found in young children, but they, unlike the schizophrenic, have not as yet developed the capacity for thought organization that the schizophrenic once had but no longer possesses. As Cameron [5, p. 59] states, " . . . the child is in the process of developing adult social language and thought organization, whereas the schizophrenic is in the process of losing it."

## SUMMARY AND CONCLUSION

1. The purpose of this study is to investigate the concept of schizophrenic regression through a study of language.

2. The percentages of nouns, pronouns, verbs and adjectives derived from the sentence completion tests were obtained from: (*a*) a schizophrenic group of 30 patients, (*b*) a comparable normal group, (*c*) groups of children at various age levels, and (*d*) a group of college students.

3. Significant parts-of-speech percentage differences were found between the schizophrenic and control group, but no significant differences were found between the schizophrenic and the fifth-grade group. As measured by the parts of speech, the language of the fifth-grade group and the schizophrenic group is similar.

4. On the basis of the parts of speech, it was concluded that the language of the schizophrenic group is "regressed" as compared with the language of a comparable normal group.

5. Interpretation of the parts of speech suggested that the schizophrenic and the child think concretistically and do not clarify their meanings in terms of a listener, either potential or real.

6. The schizophrenic group was divided on the basis of ability to organize and clarify their written language. All schizophrenics displayed disorganization and overinclusion through their speech content, but some of them displayed none of this in their sentence completion tests. It was found that disorganization and "regression" are closely related because those schizophrenics who were most disorganized also displayed "regression" by their language content. The organized schizophrenic group did not show this language "regression."

7. On the basis of the above findings, regression could be interpreted as a function of thought disorganization resulting from nonreference to, and withdrawal from, the social group. The child and the schizophrenic have similar language content because both are nonoriented with regard to consensual reality. As Sullivan [14, p. 10] states: "When one is able to relax all vigilance against one's fellows, language similarly regresses toward its beginning."

This study found that schizophrenic regression is closely related to and, perhaps, a function of thought disorganization (Table 3 and Discussion). The interpretation of the schizophrenic's behavior as a complete return to an emotional and intellectual infantilism is not supported by a qualitative analysis of the parts-of-speech percentages used by the children and schizophrenics. The theory that schizophrenic disorganization is a result of a withdrawal from consensual reality [14] should be investigated further. This study brings out no contraindicators in connection with this theory; moreover, the interpretation of the significance of the percentages of nouns and adjectives found in the schizophrenic's language would seem to support such a theory.

*Received February 5, 1951.*

## REFERENCES

1. ANDERSON, J. E. The development of motor, linguistic, and intellectual skills in young children. A report on the 3rd conference on research in child development. Toronto: Univ. of Toronto, May, 1929.
2. BALKEN, E. R., AND MASSERMAN, J. H. The language of phantasy: III. The language of the phantasies of patients with conversion hysteria, anxiety state, and obsessive-compulsive neuroses. *J. Psychol.*, 1940, **10**, 75–86.
3. BODER, D. P. Adjective-verb quotient; a con-

tribution to the psychology of language. *Psychol. Rec.*, 1940, 3, 310–343.

4. BUSEMANN, A. *Die Sprache der Jugend als Ausdruck der Entwicklungs-rhythmik.* Jena: Fisher, 1925.
5. CAMERON, N. Experimental analysis of schizophrenic thinking. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia.* Berkeley and Los Angeles: Univ. of Calif. Press, 1946. Pp. 50–64.
6. CAMERON, N. *The psychology of behavior disorders, a biosocial approach.* Boston: Houghton Mifflin, 1947.
7. ELLSWORTH, R. B. Measuring the regression of the schizophrenic patient by comparing his language with the language of the child at various age levels. Unpublished Master's thesis, Utah State Coll., 1950.
8. FAIRBANKS, HELEN. The quantitative differentiation of samples of spoken language. *Psychol. Monogr.*, 1944, 56, No. 2.
9. FENICHEL, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.
10. GOLDSTEIN, K. Methodological approach to the study of schizophrenic thought disorder. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia.* Berkeley and Los Angeles: Univ. of Calif. Press, 1946. Pp. 17-40.
11. LANDIS, C., AND BOLLES, MARJORIE. *Textbook of abnormal psychology.* New York: Macmillan, 1946.
12. MCCARTHY, DOROTHEA. *The language development of the preschool child.* Minneapolis: Univ. of Minn. Press, 1930.
13. PIAGET, J. *The language and thought of a child.* New York: Harcourt, Brace, 1926.
14. SULLIVAN, H. S. The language of schizophrenia. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia.* Berkeley and Los Angeles: U. of Calif. Press, 1946. Pp. 4–16.
15. *Webster's collegiate dictionary.* (2nd Ed.) Springfield, Mass.: G. and C. Merriam, 1946.

# EPILEPTIC INTELLIGENCE

A. LOUISE COLLINS

LYNN, MASS.

PAST studies of epileptic intelligence have seldom been based on extramural patients. Only the records of institutional or public clinic patients have been available for such study, and these often have included a large proportion of mental defectives. The intellectual defects were usually concomitants or precedent conditions and not resultants of the epilepsy, but the low intelligence found was often assumed to be that of epileptics in general [3, 4]. Normal-living private patients, who form the great majority of the country's epileptic population, conservatively estimated at a half-million [10], were not sampled. The present study is based on a sample of the records of the neglected majority, private patients who have come or have been sent by their physicians to the Seizure Unit of the Children's Medical Center in Boston for treatment under its director, Dr. William G. Lennox, neurologist and recognized authority in the field of epilepsy. An intelligence test is part of the routine examination at the Seizure Unit, the Wechsler-Bellevue test being usually chosen for adult and adolescent patients. These records, now numbering 400, have been available to the writer for evaluation and comparison with normal expectation.

## SUBJECTS

Data on age, education, and occupational status have been furnished by Dr. Lennox, as have clinical data on type and frequency of seizures, duration of the disorder, age of onset, electroencephalographic recordings, and the presence or absence of brain damage.

The age range of the group was wide, from 13 to 63 years, with a mean of 24-8. There were 228 adults aged 20 to 63, with a mean of 30-8, and 172 adolescents whose mean age was 16-5. There were 178 male and 222 female subjects.

The data for occupational status were not sufficiently specific to permit numerical ratings on the Taussig or similar scales, but they were sufficient to indicate a group well above average, with a large majority in the clerical, business, or professional classes.

The educational level of a group composed only of private patients is naturally high. Table 1 gives the ratings of the highest educational level attained, or the level at the time of testing. More that half of the group is found at the high school level, and more than one-quarter at the college, professional school, or graduate level.

TABLE 1

EDUCATIONAL LEVEL OF SUBJECTS

| Educational Level | | | *N* |
|---|---|---|---|
| Home teaching only | | | 5 |
| Special schools | | | 4 |
| Elementary grades only | | | 11 |
| Junior high school | | | 63 |
| Senior high school | | | 162 |
| College | | | 75 |
| Professional and graduate | | | 25 |
| | Law | 6 | |
| | Medicine | 9 | |
| | Theology | 5 | |
| | Graduate school | 5 | |
| Special training schools | | | 55 |
| | Hospital-nursing | 7 | |
| | Teacher-training | 3 | |
| | Vocational | 13 | |
| | Art and music | 11 | |
| | Business | 18 | |
| | Finishing school | 3 | |
| Total | | | 400 |

## IQ DISTRIBUTION OF THE WHOLE GROUP

All tests were administered according to the prescribed technique. About half were given by the writer, others were given by assisting

psychometrists[1] of the office staff of the Seizure Unit, and a few were sent in with patients by referring physicians.

TABLE 2
INTELLIGENCE CLASSIFICATION OF 400 SUBJECTS

| | IQ | N | Per Cent Obtained | Per Cent Expected |
|---|---|---|---|---|
| Defective | 65 & below | 11 | 2.75 | 2.2 |
| Borderline | 66–79 | 27 | 6.75 | 6.7 |
| Dull normal | 80–90 | 30 | 7.75 | 16.1 |
| Average | 91–110 | 128 | 32.00 | 50.0 |
| Bright normal | 111–119 | 92 | 23.00 | 16.1 |
| Superior | 120–127 | 71 | 17.25 | 6.7 |
| Very superior | 128 & over | 41 | 10.25 | 2.2 |
| Mean | 108.04 | | | |
| Median | 111.00 | | $\sigma_M = .92$ | |
| Sigma | 18.05 | | $PE_M = .62$ | |
| Range | 39–144 | | $\sigma_\sigma = .65$ | |

Table 2 gives the distribution of the obtained IQ's of the whole group, distributed according to Wechsler's classification [15]. Our distribution is not that of a normal curve, due to the large proportion of the group in the three highest grades. The normal expectancy percentages are appended for comparison.

If the scores are divided according to sex, the female average, 108.63, is slightly above the male, 107.69, but the reverse is true of the median—male, 111.15, and female, 110.5. Variability, shown by the standard deviation, is greater in the male sample.

In comparing the verbal and performance IQ's, we found 10 more subjects had a higher performance IQ than had a higher verbal IQ (196 and 186, 18 the same). Wechsler considers a 10-point difference to be within normal range. Sixty-eight of our subjects scored more than 10 points higher on the verbal scale, 64 on the performance.

## PATTERN ANALYSIS OF THE WHOLE GROUP

For a detailed analysis of the test results, Table 3 gives the relative ranking of the subtest scores of the 400 subjects with means and standard deviations. Deviation scores have also been computed, i.e., each subject's deviation on each subtest from his own subtest average. The resulting rank order of the deviation scores is the same as that of the subtest averages.

Analyzing these scores, we find the reason for the inferiority noted in the verbal scale. The two lowest-ranking subtests are Arithmetic and Digit Span, which Wechsler [15] includes in his verbal scale. If we follow Rapaport's suggestion [13] as to rationale and consider Digit Span and Arithmetic as concentration and not verbal tests, we should find a distinct verbal superiority in our cases. Our means would be—verbalized tests, 11.31; performance tests, 10.55; and concentration tests, 9.64.

The subtest records of the present study fall quite definitely into a pattern of three distinct groups, a high-ranking group comprising the four verbalized tests and Block Design, an intermediate group composed of the other four performance tests, and the low-scoring group, Digit Span and Arithmetic. Very little divergence from this general pattern has been found as the whole group has been divided in various ways according to clinical factors. Such changes as have appeared in the three group rankings are usually intragroup and not intergroup changes.

Comparing records of the two sexes, we found differences insignificant in 8 of the 11 averages, significant in 3 at the 1 per cent level of confidence. These three were the male Arithmetic, $t = 3.31$, male Information, $t = 3.72$, and the female Digit Symbol, the very high $t = 5.5$.

TABLE 3
SUBTEST SCORES OF 400 PRIVATE PATIENTS

| Rank | Test | Mean | Sigma |
|---|---|---|---|
| 1 | Comprehension | 11.88 | 3.38 |
| 2 | Similarities | 11.30 | 3.59 |
| 3 | Block Design | 11.21 | 3.61 |
| 4 | Vocabulary | 11.08 | 2.97 |
| 5 | Information | 10.98 | 3.19 |
| 6 | Picture Completion | 10.69 | 3.06 |
| 7 | Object Assembly | 10.56 | 3.61 |
| 8 | Picture Arrangement | 10.19 | 3.40 |
| 9 | Digit Symbol | 10.10 | 3.15 |
| 10 | Arithmetic | 9.66 | 3.85 |
| 11 | Digit Span | 9.17 | 3.32 |

[1]Especially should acknowledgments be made to Mrs. Pauline Hahn and Mrs. Marjorie Archibald, who administered 150 of the tests.

### ONSET AGE AND DURATION

Early onset of the disorder and long duration have long been considered as detrimental factors. Using our clinical data on these points, we have computed the correlations with the obtained IQ's by the Pearson product-moment method. The results are as follows:

Onset age and IQ, $r = +.25$, $PE = .026$
Duration and IQ, $r = -.06$, $PE = .032$

The onset age coefficient, over nine times its probable error, is significant at the 1 per cent level of confidence; the duration coefficient is not significant even at the 10 per cent level. Thus, the effect of early onset would appear to be decidedly detrimental, and the effect of long duration far less so.

In interpreting the onset-age figure, it should be noted that it concerns only present psychometric results, not original endowment, and implies nothing as to whether the cause of the suggested deficit is an organic effect of the disorder that retards normal development, or whether it is a social and environmental effect.

### AGE GROUP COMPARISON WITH STANDARDIZATION NORMALS

Comparison has been made with the Wechsler-Bellevue standardization group [15] by dividing our cases into similar age groups. Our adolescent 17- to 19-year-old group contains practically the same number as the standards; each of the three youngest adult groups (20-24, 25-29, 30-34) numbers more than 25 per cent of the corresponding standards; the other younger and older groups are smaller, and comparison is probably less valid. The records of all these groups have been tabulated and compared. Without reproducing here our many charts, we cite the more important of the findings derived from them.

All the mean IQs of the patient age-groups are above those of the standards, the means of the older groups showing the greatest superiority. Variability was greater in three of the standard groups; in all others, the patient group was more variable.

In the pattern analysis, rank order not only differs much from the standards, but it is more consistent from age to age. Performance tests are often among the leaders in the standards, seldom in the patients, except for Block Design. The verbalized tests show the greatest increment above the standards.

### PATTERN ANALYSIS OF THE AGE GROUPS

The subtest patterns of the patient groups of various ages conform fairly closely, but in no case exactly, to the general pattern of the 400 subjects. The following characteristics were noted:

A few performance records are among the leaders in the youngest groups.

Comprehension always ranks high and leads in all the larger and older groups.

Block Design is always among the first five.

Verbalized tests are with very few exceptions in the first five, never in the low group.

Digit Span and Arithmetic are always low; Arithmetic rises slightly in position in the older groups.

Similarities varies most in rank; Vocabulary, least.

The most striking finding would seem to be the marked superiority of the older groups.

### CLASSIFICATION AS TO TYPES OF SEIZURE

More than four-fifths of our 400 subjects have suffered grand mal seizures; of these half have had the grand mal type only, and half have had either petit mal or psychomotor seizures in addition to the grand mal. Of the remaining one-fifth, 31 have had only psychomotor seizures, 26 only petit mal, and 8 subjects, all three types.

Classifying the IQ's according to these types of seizure, we find that the groups which contain subjects whose seizures are wholly or partly of the petit mal type rank highest, the psychomotor cases next, and the grand mal subjects a little lower. Differences, however, are not great, and all three groups average above the population norm.

The grand mal subjects, because of their great preponderance of numbers, are the major determinant of the subtest averages of the whole 400. The rank order of their subtest records shows the same general grouping of high, medium, and low with few and slight changes. The petit mal cases also conform equally closely to the general pattern.

The variations in the pattern found were

largely in the psychomotor groups; the chief feature of this change was for Arithmetic, which moved up from the low to the middle group, and Digit Symbol and Picture Arrangement, which moved down to the low.

An estimate of these findings, especially the lower grand mal records, should take account, not only of the records, but of the environmental factors, particularly the probability that the grand mal type of the disorder would interfere more than the others with normal social and educational opportunities.

### ELECTROENCEPHALOGRAPHIC CLASSIFICATION

For comparison of electrocephalographic recordings with intelligence ratings, the 400 patients have been divided according to the Gibbs classification [6], as follows:

Normal
Mildly slow or fast ($S_1$ $F_1$)
3/sec dart & dome (petit mal, PM)
Very slow or fast ($S_2$ $F_2$)
2/sec spike and wave (petit mal variant, PMV)
Other (miscellaneous)

The first three groups show little or no dysrhythmia; the next two are very abnormal with much dysrhythmia; the last is a mixed group of cases not readily classifiable, and includes some psychomotor-type brain waves (HVS).

Table 4, which gives the IQ averages of these groups, indicates the superiority of the subjects with the most normal brain waves, the intermediate position of the miscellaneous group, and the much lower ranking of the groups showing the most dysrhythmia.

The critical ratios of the differences between the groups (Table 5) indicate those differences that are significant at the 1 per cent level of confidence. Differences between any two of the more normal groups are not significant, nor are those between the two abnormal groups.

In the pattern analysis, rank order conformed much more to the general pattern and showed the least variability in the more normal groups, but varied much more in the two abnormal groups; in these latter groups, there were some instances of performance tests ranking among the leaders, especially in the PMV group.

TABLE 4
CLASSIFICATION OF IQ'S ACCORDING TO ELECTROENCEPHALOGRAPHIC RECORDINGS

| | *N* | Mean IQ | Sigma |
|---|---|---|---|
| Normal | 46 | 113.00 | 12.75 |
| $S_1$ $F_1$ | 143 | 109.30 | 19.00 |
| PM | 98 | 110.40 | 15.60 |
| Misc. | 63 | 108.08 | 16.45 |
| PMV | 20 | 93.25 | 18.55 |
| $S_2$ $F_2$ | 30 | 93.33 | 24.55 |

TABLE 5
CRITICAL RATIOS OF THE IQ DIFFERENCES BETWEEN GROUPS CLASSIFIED BY ELECTROENCEPHALOGRAPHIC RECORDINGS

| | $S_1$ $F_1$ | PM | Misc. | PMV | $S_2$ $F_2$ |
|---|---|---|---|---|---|
| Normal | 1.50 | 1.08 | 1.76 | 3.89** | 3.98** |
| $S_1$ $F_1$ | | .46 | 0.47 | 3.24** | 3.23** |
| PM | | | 0.89 | 3.53** | 3.53** |
| Misc. | | | | 2.88** | 2.94** |
| PMV | | | | | 0.01 |

**Significant at the 1 per cent level of confidence.

### TRAUMATIC AND ESSENTIAL CASES[2]

*IQ Analysis.* Early writers usually maintained that the constitutional, "idiopathic," type of epilepsy had a more deteriorating effect on the patient's mentality than did that due to brain injury. Our cases have been divided according to the clinical report on this factor, into 299 essential and 101 traumatic subjects, and the averages of both groups have been compared for IQ's and for subtest scores, and the critical ratios have been computed. The result definitely contraindicates the early belief, for the essential subjects average more than 12 IQ points above the traumatic. The essential average is 111.20; the traumatic, 98.95. Below 91 IQ, the lower limit of Wechsler's normal group, there were 31 per cent of the traumatic records and 11 per cent of the essential. The critical ratio of the difference in means is 5.4, significant at the 1 per cent level.

When the verbal and performance scales were considered separately, the results were similar, the average difference being a little

[2]The term "traumatic" has not been used here in its most limited technical sense, but refers to any form of brain damage. The term "essential" is used in place of the old term "idiopathic" for the genetic or physiological cases.

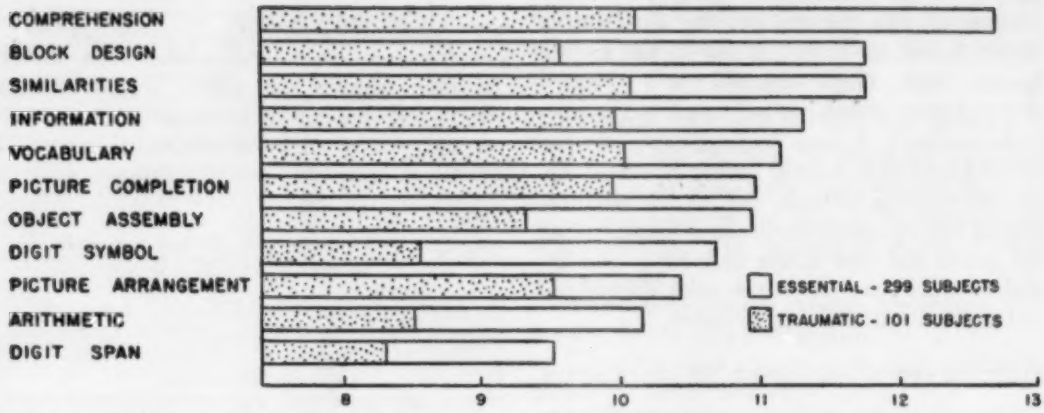


FIG. 1. Subtest score averages of the traumatic and essential cases.

TABLE 6

RANK ORDER OF SUBTEST AVERAGES FOR TRAUMATIC AND ESSENTIAL CASES

| Traumatic Cases | Mean | Rank | Essential Cases | Mean |
|---|---|---|---|---|
| Comprehension | 10.48 | 1 | Comprehension | 12.64 |
| Similarities | 10.02 | 2 | Block Design | 11.77 |
| Vocabulary | 10.00 | 3 | Similarities | 11.76 |
| Information | 9.95 | 4 | Information | 11.33 |
| Picture Completion | 9.94 | 5 | Vocabulary | 11.18 |
| Block Design | 9.52 | 6 | Picture Completion | 10.97 |
| Picture Arrangement | 9.46 | 7 | Object Assembly | 10.96 |
| Object Assembly | 9.16 | 8 | Digit Symbol | 10.65 |
| Digit Symbol | 8.47 | 9 | Picture Arrangement | 10.40 |
| Digit Span | 8.27 | 10 | Arithmetic | 10.14 |
| Arithmetic | 8.26 | 11 | Digit Span | 9.49 |

larger for the performance scale, but for both scales the differences are significant at the 1 per cent level.

Variability ($\sigma$) in the full scale and in both verbal and performance scales was greater in the traumatic group. Of the 50 subjects reported in the previous section whose electroencephalographic records showed marked dysrhythmia, 32 were of the traumatic type.

*Pattern Analysis.* When the scores of the individual subtests were considered, marked differences were found in the averages, all in favor of the constitutional subjects. Table 6 shows the comparative rank order of the two lists, and Figure 1 indicates the relative rating of the two groups on each test. The most interesting feature of the difference in rank in the two lists is the drop of Block Design from second place in the genetic subjects to sixth in the traumatic list. Similar findings as to traumatic inferiority on this type of test have been reported in previous studies of epilepsy [16] and other forms of brain disorder [7, 8].

All critical ratios of the differences in the subtest averages of the two groups are significant at the 1 per cent level except for Picture Arrangement, which is significant at the 2 per cent level. The $t$ scores are largest for the verbal and Block Design tests, with one exception; Digit Symbol, the chief speed test, yields the largest $t = 5.93$.

Variability, as shown by the standard deviations, was found to be greater in the traumatic list for all but two subtests—Arithmetic, the lowest of the traumatic averages and Picture Arrangement, relatively higher in the traumatic than in the essential list.

### SUBTEST PATTERNS OF GROUPS CLASSIFIED ACCORDING TO IQ

Since our subjects rank definitely above the population average in intelligence, the question naturally arises as to whether the ability pattern which we have found in the records of the group—both the group as a whole and,

to a great extent, however clinically divided—is characteristic of the disorder, or is due to the fact that we are dealing with a group of better than average intelligence and education.

To test this, we classified our cases into four groups according to IQ—Group I, 113 superior subjects, IQ above 119; Group II, 91 above-average subjects, IQ 111 to 119; Group III, 127 average subjects, IQ 91 to 110; and Group IV, 69 below-average subjects, IQ below 90. The deviation scores showing relative, not actual, subtest ratings were used in this comparison. The most striking findings are as follows:

Deviations ranged from +1.48 for Comprehension in Group I, to -1.92 for Digit Span in Group II and Arithmetic in Group III.

The greatest range of deviations was found in Group II, the above-average subjects, and the least in the low scorers.

While all four groups approximate the observed general pattern, the conformity is much closer in the higher-ranking groups.

Performance tests begin to appear in somewhat higher positions in the average group.

Comprehension ranks first in all groups.

Block Design drops in rank, with lower IQ's, from second place in the superior group to a small minus deviation in the low group.

Similarities also steadily decreases with decreasing IQ.

Object Assembly and Digit Symbol rank relatively higher in the lower groups.

Arithmetic rises from lowest position, with the largest minus deviation, in the lower groups, to sixth place, with near-zero deviation in the superior group.

Digit Span, failure on which is often considered an anxiety indicator [13], is always one of the two lowest ranking tests.

From these findings we might suggest that our general pattern—indicating superiority in thinking, reasoning power, and concept formation—may be due to the superior intelligence and education of our group, but that the consistently high rating of Comprehension in all our subgroups and the equally consistent low rating of Digit Span, with its anxiety implication, might be justifiably considered as an epileptic characteristic.

## DETERIORATION THEORIES

Several theories have considered evidence of deterioration furnished by differential ability ratings: chiefly, that of Babcock that vocabulary holds up in age and psychosis when other functions fail [1]; that of Bovet [2] and Simmins [14], who posit verbal inadequacy in the epileptic; that of Wechsler who proposes "hold" and "don't hold" tests [15]; and that of Rapaport, who postulates impairment when Comprehension is below Information [13].

Vocabulary scatter has been considered for our group as a whole and as divided into the genetic and traumatic subgroups. Vocabulary ranks as the lowest of the verbalized tests for the whole group and for the essential cases; for the traumatic, it ranks third, above Information. Indication of deterioration would seem slight here, and only in the traumatic cases. The position of Vocabulary above all but the blocks of the performance scale seems to contraindicate the verbal inadequacy theory. Comprehension is above Information in all groups. By the Wechsler method, our high Block Design usually balances the low Digit Span in the "don't hold" tests, so that few of the many subgroups show over 10 per cent deterioration, and none over 20 per cent, his criterion of significance. While a few individuals have been found to show definite deterioration, there is little evidence of it as a group characteristic.

## SUMMARY

Our study has dealt with the results obtained on the Wechsler-Bellevue intelligence tests of 400 adult and adolescent epileptic, extramural, office patients. It is an unselected group, not one chosen for a specific experiment. Being unselected, the range is great in age, education, intelligence, and type and severity of seizures. Therefore, the records have been studied, not only as a whole, but with the group classified in various ways according to clinical records.

Our findings are:

The group as a whole shows a wide range in obtained IQ's, but averages well above normal expectancy.

We have found little difference in the full IQ's of the two sexes. There are, however, significant deviations in Arithmetic and

Information scores in favor of the male subjects, and in Digit Symbol in favor of the female.

Comparison of our records with those of Wechsler's standardization normals shows generally higher averages and a more consistent pattern of subtests, favoring our group, especially our older subjects.

There is little evidence of deterioration except in the organic cases; in fact, the older subjects show consistently higher records, not only in IQ, in which there is allowance for age, but also in the subtest scores with no such allowance.

We have found a positive correlation of IQ with onset age, and a very low negative correlation of IQ with the duration of the disorder.

The most definite and probably most significant of our findings is the superiority of the constitutional group over the subjects whose disorder is due to brain damage.

We have found a fairly distinct pattern of subtest records that characterizes all of our many groups, with only minor variations. This pattern consists in general, of three distinct groupings: (1) the four verbalized tests and Block Design, (2) the other four performance tests, and (3) the low-rating Digit Span and Arithmetic. There is some intragroup variability in these groupings in our various classifications, but few intergroup shifts. The few variations from this pattern were chiefly in the traumatic cases, those whose seizures were of the psychomotor type, the groups of lowest intelligence, and those with the most marked dysrhythmia. Comprehension is the highest-ranking subtest in all groups. A feature of the traumatic-genetic comparison is the significant drop of Block Design in the brain-injured cases.

Comparison of our results with those of similar investigations of other types of disorder have indicated no conformity with observed patterns of psychotics or pyschopaths, and little similarity to those of neurotics. Our results also differ materially from those found in investigations of deterioration [1, 9, 11, 12, 15].

## CONCLUSIONS

Recognizing and allowing for the fact that our subjects, as a private patient group, are above the population norm in intelligence and education, and that some of our reported findings are due to this superiority, we offer the following conclusions as suggested by the study:

The noted superiority may be taken as evidence of the fact that mental defect or low intelligence is not necessarily a concomitant or a resultant of epilepsy. The disorder may occur in individuals of all degrees of intelligence.

Epilepsy due to brain injury has a more deleterious mental effect than does the genetic form. It is in the former, not the latter type, that the deterioration, sometimes attributed wholly to the epilepsy, may usually be noted.

Anxiety, naturally engendered by the liability to seizures, might be a determinant of some features of the rather consistently observed pattern of response in our group.

Environmental causes, particularly those that operate to prevent normal educational opportunities and normal social contacts, may be a potent factor in causing the apparent deterioration, formerly considered inevitable. Possibly drug effects of the older types of therapy might be numbered among the environmental causes.

Finally, we suggest the inadequacy of some of the old-time theories of the disorder, based only on institutional cases and completely neglecting the large normal epileptic population.

*Received February 12, 1951.*

## REFERENCES

1. BABCOCK, HARRIET. An experiment in the measurement of mental deterioration. *Arch. Psychol., N.Y.*, 1930. No. 117.
2. BOVET, T. Der Rorschachversuch bei verscheidenen Formen von Epilepsie. *Arch. Neurol. Psychiat., Genève*, 1936, **37**, 156–157.
3. COLLINS, A. LOUISE. Psychometric records of institutionalized epileptics. *J. Psychol.*, 1941, **11**, 359–370.
4. COLLINS, A. LOUISE, ATWELL, C. R., AND MOORE, M. Stanford-Binet response patterns in epileptics. *Amer. J. Orthopsychiat.*, 1938, **8**, 51–63.
5. GARRETT, H. E. *Statistics in psychology and education.* (3rd Ed.) New York: Longmans, Green, 1947.
6. GIBBS, F. A., GIBBS, E. L., AND LENNOX, W. G.

Electroencephalographic classification of epileptics and controls. *Arch. Neurol. Psychiat., Chicago,* 1943, **50**, 111.

7. Goldman, Rosaline, Greenblatt, M., and Coon, G. P. Use of the Bellevue-Wechsler Scale in clinical psychiatry. *J. nerv. ment. Dis.*, 1946, **104**, 144–179.
8. Greenblatt, M., Levin, S., and Atwell, C. R. Comparative value of electro-encephalogram and abstraction tests in diagnosis of brain damage. *J. nerv. ment. Dis.*, 1945, **102**, 383–391.
9. Johnson, Anna P. Measuring mental deterioration by the differential test score method. *Amer. J. ment. Def.*, 1947, **51**, 389–390.
10. Lennox, W. G. *Modern therapy in general practice.* Baltimore: Williams & Wilkins, 1940.
11. Magaret, Ann. Parallels in the behavior of schizophrenics, paretics and pre-senile non-psychotic patients. *J. abnorm. soc. Psychol.*, 1942, **37**, 511–526.
12. Rabin, A. I. Differential psychometric patterns in schizophrenia and manic-depressive psychosis. *J. abnorm. soc. Psychol.*, 1942, **37**, 270–272.
13. Rapaport, D., Gill, M., and Schafer, R. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1945. Vol. I.
14. Simmins, Constance. Mental incapacity: The intelligence of patients in mental hospitals. *Character & Pers.*, 1935, **4**, 25–33.
15. Wechsler, D. *The measurement of adult intelligence.* (3rd Ed.) Baltimore: Williams & Wilkins, 1944.
16. Winfield, D. L. Intellectual performance of cryptogenic epileptics, symptomatic epileptics and posttraumatic encephalopaths who were members of the armed services during World War II. *J. abnorm. soc. Psychol.*, 1951, **46**, 336-343.

# INDUCED FRUSTRATION AND PUNISHMENT-REWARD EXPECTANCY IN THEMATIC APPERCEPTION STORIES[1]

VAUGHN J. CRANDALL
THE FELS RESEARCH INSTITUTE FOR THE STUDY OF HUMAN DEVELOPMENT

THE adjustment of an individual to his environment may be assessed in many ways. Usually, this has been done either by obtaining a general picture of the individual's personality make-up or by evaluating his present reactions to particular situations. The TAT has been used for both purposes. While the TAT has been more frequently used for global personality assessments, several investigators have employed it to evaluate frustration reactions with promising results [3, 12, 13, 14, 15].

The present investigation was concerned with the evaluation of individuals' reactions to a specific situation: induced frustration. More specifically it dealt with the relationship between induced frustration and resultant changes in thematic apperception stories.

The experimental variable of the present study was induced frustration concerning physical skills prowess. The dependent variable was punishment-reward expectancy inferred from the subjects' thematic apperception stories. Strong reward-expectancy was inferred when the reinforcing agents in the individual's environment were perceived by him as helpful, friendly, nonthreatening, etc. Strong reward-expectancy was also inferred when the individual's behavior was primarily oriented toward attaining the goal rather than defending himself against possible failure. In contrast, strong punishment-expectancy was inferred when the individual interpreted the presses of his environment as hostile, rejecting, impeding, etc., and when his behavior was primarily oriented toward defending himself against failure. Punishment-reward expectancy was defined as "the reinforcement value of relevant social reinforcing agents for the particular behavior employed by the individual in his strivings to satisfy a given need."[2] Specific behavioral referents in thematic apperception stories for the concept punishment-reward expectancy will be given in a subsequent discussion of the development of the scoring manual.

It was hypothesized that frustration would be followed by increased punishment expectancy and that this would be reflected in the subjects' thematic apperception stories. Stated in the null form, the hypothesis was: Punishment-reward expectancy reflected in thematic apperception stories told by subjects following frustration will not differ from punishment-reward expectancy in stories of nonfrustrated subjects, beyond that which might occur by chance.

## PROCEDURES

One hundred and fifty-eight subjects were used in various aspects of the study. They were all white male college students, members of the introductory psychology course at Ohio State University. The age range was roughly from 18 to 35. To insure a random selection of these students, their names were drawn from an alphabetized list.

The thematic apperception pictures used were developed specifically for the present

[1]Adapted from a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology, Ohio State University. The author wishes to express appreciation to Professors Julian Rotter, George A. Kelly, and Delos D. Wickens for their helpful suggestions, and to Mr. Erwin J. Lotsof, Mr. Richard Jessor, and Mr. Alvin A. Lasko who served as raters. Mr. Alfred Castaneda conducted the frustration situation.

[2]This concept is somewhat similar to the concepts of "freedom of movement" and "availability of behavior" developed by Rotter in his Social Learning Theory of Personality [17, 18].

investigation. Criteria were established for their selection. Of the following criteria the first three were adapted from a list by Heppell [8]; the rest were developed by the present writer:

1. The picture situations should be related to the particular problem of the investigation.
2. Pictures should represent situations familiar to the subjects.
3. The pictures should contain a central figure of the same sex and approximately the same age as the subject.
4. There should be a maximum number of pictures to insure reliability consistent with a minimum number to maintain high motivation in the subject.
5. The pictures should contain situations relevant to important needs and strivings of the subject. They should represent both situations entailing strivings similar to those frustrated in the experiment and ones eliciting strivings other than these.[3]
6. Pictures representing situations eliciting a given need should call forth stories eliciting that need in a clear majority of the subjects.
7. Two roughly equivalent sets of pictures should be developed. This is necessary in order to match the experimental and control subjects on the dependent variable in a pretest before the introduction of the experimental variable.

Using these criteria, a large group of pictures was gathered from which judges selected two tentative sets. These pictures were tried out in group administration to 78 subjects. Pictures not meeting the criteria were eliminated, and a final group of 18 pictures, 9 in each set, was thus obtained. Each set had 3 pictures found to elicit stories dealing with physical skills recognition, 3 pictures concerning academic recognition, and 3 pictures calling forth stories about affection toward opposite-sex peers.

A system for scoring punishment-reward expectancy in the stories was developed. Many of the scoring systems suggested in the literature were intended for a general assessment of the individual and were thus not applicable to the present investigation. However certain aspects of several scoring systems [1, 2, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 19] appeared relevant as potential indicators of punishment-reward expectancy in the present study. These potential indicators were sorted to minimize overlap and placed in one of the four following categories: environmental press of the story, adequacy of the central character's behavior in the story, formal characteristics of the story, and the subject's reactions to the story-telling situation.

A pilot study was run in which the pictures were administered to ten experimentally frustrated subjects and ten nonfrustrated subjects in individual sessions. Their stories were analyzed for the occurrence of the indicators, and certain tentative evaluations were made:

1. The indicators of the category, *environmental press*, were found to discriminate between frustrated and nonfrustrated subjects. Stories told by frustrated subjects contained more threatening, hostile, unpredictable, and obstructive environmental presses.
2. All indicators of the category, *adequacy of the central character's behavior*, discriminated. Indicators called failure-oriented behaviors appeared much more frequently following frustration. These were: withdrawal, anger and hostility, guilt, extreme suspiciousness, extreme dependence on others, daydreaming not followed by constructive action, and reality denial. Other indicators of inadequacy of the central character's behavior increased following frustration but not as markedly as failure-oriented behaviors. These included instances in which (*a*) the central character questioned the worth of the goal toward which he was striving, (*b*) needs other than the dominant need of the story conflicted with this need, and (*c*) the central character expressed doubts concerning his adequacy.
3. The only useful indicator in the category, *formal characteristics of the story*, was unrealistic endings. Endings were defined as unrealistic when they could not be predicted from the body of the story.
4. The only aspect of the *subject's reaction to the story-telling situation* which occurred more frequently following frustration was inability to develop a plot, coupled with expressed feelings of inadequacy concerning this. This took place when the subject gave little or no description or plot and verbalized a feeling of inadequacy concerning it.

A judges' scoring manual was developed using those indicators which appeared promising. It may be noted that the indicators listed were punishment-expectancy indicators. Reward-expectancy indicators were seen in stories having descriptions of the environment

[3] The suggestion that pictures should be selected which elicit strivings other than that striving which is frustrated is particularly relevant when the question of specificity of reinforcement is raised. That aspect of this investigation will be covered in a subsequent article.

or central character's behavior which used adjectival opposites of these punishment-expectancy indicators.

The manual contained a discussion of all indicators and examples of each. The judges were asked to assess the intensity and duration of all indicators in a given story before assigning a rating to that story. The ratings ran from 1 to 7, with 1 indicating strong reward-expectancy and 7 indicating strong punishment-expectancy. The judges were asked to compare the story being rated with the manual, which contained a rated story example for each rating 1 through 7 for each of the 18 pictures used.

Sixty new subjects randomly assigned to the control or experimental groups were used in the experiment proper. Each subject in the experimental group told stories to one set of pictures, underwent the induced frustration, and subsequently told stories to the other set. Each subject in the control group followed a similar procedure, except a rest period comparable in length to the time taken to administer the frustrating situation took the place of the frustration. Both the story-telling and the induced frustration situations were conducted individually.

After the first story-telling situation, each experimental subject was asked to take part in what was described as another experiment. This procedure was adopted so that the subject would not perceive the frustration and the story-telling as part of the same experiment. There was no evidence that any subject did connect the two. The writer conducted the story-telling sessions; an assistant handled the frustration situation.

The subject was introduced to the frustrator, who told him he would be asked to perform certain physical skills tasks which predicted general athletic ability. The frustrator stated that information had already been obtained on a large group of Ohio State University men and suggested that the subject and he could compare the subject's performance with these norms. The norms were fictitious and set extremely high so that it was impossible to score above the twentieth percentile on any of the six tasks. After each task, the subject's performance was compared to the norms. The frustrator's role was one of "sympathetic encouragement in the face of the subject's inadequate performance." Following the frustration, the subject immediately returned to the story-telling room and told stories to the second set of pictures.

The two sets of thematic apperception pictures were counterbalanced during administration and the pictures within each set randomly serialized.

All stories were recorded on Gray Audograph disks and subsequently transcribed. Each story was typed on a separate piece of paper and contained no identifying material concerning the subject. The experimenter and two other judges independently rated all 1,080 stories of the 30 experimental and 30 control subjects. All stories concerning a given picture were rated in a group.

## RESULTS

Two aspects of the present investigation call for evaluation. The communicability and objectivity of the rating system developed for the variable punishment-reward expectancy must be assessed. The hypothesis that frustration is followed by increased punishment expectancy must be checked.

An estimate of interjudge reliabilities was obtained in the following manner: Each subject was given a score for stories told in the first story-telling situation, by adding the ratings of his nine stories told at that time. The subject was assigned a score in a similar fashion for the second story-telling situation. The interjudge reliability was evaluated by correlating these scores obtained from each of the three judges' ratings. The correlation (Pearson $r$) of Judge A (the experimenter) with Judge B was .75, of Judge A with Judge C was .71, and of Judge B with Judge C was .67.

The experimental hypothesis was tested, using the experimenter's ratings, in the following manner: Experimental and control subjects were individually matched on the basis of their scores for the first story-telling situation. Mean gains, the difference in punishment expectancy from the first to second story-telling situation, were computed for the experimental and control subjects. Punishment expectancy of the experimental subjects exceeded that of control subjects by a mean

difference of 3.703, significant at the .001 level of confidence.

The results suggest that the system developed for rating punishment-reward expectancy was reasonably communicable. They also suggest that, within the limits of this experiment, the subjection of individuals to a frustrating situation is followed by an increased punishment-expectancy reflected in thematic apperception stories told subsequently.

## SUMMARY

The present study was designed to investigate the relationship between induced frustration and punishment-reward expectancy reflected in thematic apperception stories. Two roughly equivalent sets of thematic apperception pictures were specifically developed for this experiment according to stated criteria. Potential indicators of punishment-reward expectancy in thematic apperception stories were gleaned from the existing literature. The utility of these indicators was checked in a pilot study. Results of this preliminary study suggested that most formal characteristics of the thematic apperception stories did not differentiate frustrated from nonfrustrated subjects. Neither did various reactions of the subjects to the story-telling situation. However, when the rater attended to the adequacy of the central character's behavior in the story and to whether the environmental presses were threatening or nonthreatening, it was possible to differentiate the stories of subjects who had been frustrated from those who had not.

A scoring system was developed using these indicators and was applied to the stories of a new group of subjects in the experiment proper. The magnitude of the interjudge correlations suggested that this scoring system was reasonably communicable and objective.

In the experiment proper an experimental group of 30 subjects was frustrated; a control group of 30 subjects was not. A statistically significant increased punishment expectancy was found in the stories of the frustrated subjects.

*Received January 19, 1951.*

## REFERENCES

1. ARON, BETTY. *A manual for analysis of the Thematic Apperception Test: A method and technique for personality research.* Berkeley, Calif.: Willis E. Berg, 1949.
2. ATKINSON, J. W., AND MCCLELLAND, D. C. The projective expression of needs. II. The effect of different intensities of the hunger drive on thematic apperception. *J. exp. Psychol.*, 1948, **38**, 643–658.
3. BELLAK, L. An experimental investigation of projection. *Psychol. Bull.*, 1942, **39**, 489–490.
4. CLARK, R. W. A method of administering and evaluating the Thematic Apperception Test. *Genet. Psychol. Monogr.*, 1944, **30**, 5–53.
5. FRENKEL-BRUNSWIK, ELSE. Dynamic and cognitive categorization of qualitative material: I. General problems and the Thematic Apperception Test. *J. Psychol.*, 1948, **25**, 253–260.
6. GARFIELD, S. L., AND ERON, L. D. Interpreting mood and activity in TAT stories. *J. abnorm. soc. Psychol.*, 1948, **43**, 338–345.
7. HENRY, W. E. The Thematic Apperception Technique in the study of culture-personality relations. *Genet. Psychol. Monogr.*, 1947, **37**, 3–135.
8. HEPPELL, H. K. Selecting pictures to obtain fantasy material on a specific topic. Unpublished Master's thesis, Ohio State Univ., 1948.
9. LOWELL, E. L., AND ATKINSON, J. W. Scoring projective records for need for achievement (1st draft), 1950. (Mimeo.)
10. MAYMAN, M., AND KUTNER, B. Reliability in analyzing Thematic Apperception Test stories. *J. abnorm. soc. Psychol.*, 1947, **42**, 365–368.
11. MCCLELLAND, D. C., ATKINSON, J. W., AND CLARK, R. A. The effect of different intensities of need mastery on thematic apperception. ONR report, 1947. (Mimeo.)
12. MCCLELLAND, D. C., ATKINSON, J. W., AND CLARK, R. A. The projective expression of needs. III. The effect of ego involvement, success and failure on perception. *J. Psychol.*, 1949, **27**, 311–330.
13. MCCLELLAND, D. C., CLARK, R. A., ROBY, T. B., AND ATKINSON, J. W. Need for achievement on thematic apperception. *J. exp. Psychol.*, 1947, **39**, 242–259.
14. MCCLELLAND, D. C., CLARK, R. A., ROBY, T. B., AND ATKINSON, J. W. The projective expression of needs. IV. The effect of the need for achievement on thematic apperception. *J. exp. Psychol.*, 1949, **39**, 242–255.
15. RODNICK, E. H., AND KLEBANOFF, S. G. Projective reactions to induced frustration as a measure of social adjustment. *Psychol. Bull.*, 1942, **39**, 489. (Abstract.)
16. ROTTER, J. B. Studies in the use and validation of the Thematic Apperception Test with

mentally disordered patients. I. Method of analysis and clinical problems. *Character & Pers.*, 1940, 9, 18–34.

17. ROTTER, J. B. Tentative formulation of basic principles for a social learning theory of personality, 1948. (Mimeo.)
18. ROTTER, J. B. Tentative formulation of some social learning constructs, 1950. (Mimeo.)
19. SHORR, J. E. A proposed system for scoring the TAT. *J. clin. Psychol.*, 1948, 4, 189–194.
20. WYATT, F. The scoring and analysis of the Thematic Apperception Test. *J. Psychol.*, 1947, 24, 319–330.

# PROJECTIVE PICTURES AS INTERVIEW DEVICES

H. K. HEPPELL
SAN DIEGO CITY SCHOOLS
AND
V. C. RAIMY
UNIVERSITY OF COLORADO[1]

THE diagnostic interview is essentially a method for obtaining information from child or adult upon which clinical judgments as to etiology, diagnosis, or prognosis can be based. Without minimizing the necessity and usefulness of the conventional question-answer-discussion type of interview, one can point to the possibility of obtaining helpful clinical material by methods which depart from tradition yet serve the basic purpose. There is no need to provide additional testimonials to projective pictures such as the TAT cards. As *survey* techniques they are part of standard clinical procedure.

However, situations often arise in which the psychologist or his clinical colleague wishes additional information on specific topics for which the standard sets of pictures provide little stimulus material or for which the completed protocol reveals conflicting or ambiguous data. The TAT can thus be viewed as a set of picture situations designed to sink shafts at various points in the subject's apperceptive or motivational mass with the reasonable hope that major trends will be revealed. Nonetheless, even the most ardent supporters of this technique would be unlikely to argue that the sinking of shafts provides all the information which may be needed by the clinician or that inadequate sampling will not sometimes result. Moreover, with a trend or theme of suspected significance once revealed, further exploration and elaboration is often required but at present can be obtained only by ingenious or ingenuous guesses from "related protocol" or by resorting to ordinary interview procedures.

Concretely, there are frequent occasions in work with children or adults when more information is badly needed on the subject's perception of his sexual role, his family role, his social relationships with peers, his scholastic relationships, etc. In addition to clinical work, many research problems would tap regions enormously rich in psychological material if reasonably valid techniques were available for exploring specified aspects of the individual's self-concept in its personal and social integrations.

The present article makes no attempt to examine the validity of data obtained from pictures designed to elicit stories for a specified topic, since validation of projective stories presents major theoretical problems which are outside our scope at the present time. The attempt has been made, however, to extend the use of the projective picture *technique* into the area of specific topics. Contributory literature seems meager. Rapaport [4] has proposed several TAT "area" cards, but one finds very few pictures applicable to each area. Rosenzweig and coworkers have approached the problem in "Picture-Frustration Studies" [5, 6, 7] but with emphasis almost entirely on the picture technique as a method for eliciting data on frustration and its place in personality patterns. The problem of the present study is similar except that we have been interested primarily in developing a method for selecting "good pictures" for any given topic and within a given age or sex group.

## METHOD

For reasons of convenience, the study was

[1]This study was carried out at Ohio State University and the State of Ohio, Bureau of Juvenile Research. The authors wish to thank Dr. Persis Simmons and Dr. Maurice Newburger for their help.

concerned with child-parent relationships and arbitrarily limited to boys between the ages of 10 and 13 years institutionalized at the State of Ohio, Bureau of Juvenile Research.

Fifty cartoon pictures which seemed to provide possibilities for eliciting stories of child-parent relationships were selected originally by inspection. Cartoons were used without proof of their greater suitability for provoking imaginative productions. It was assumed, however, that they would provide a medium attractive to children. Furthermore, cartoons usually contain a minimum of detail and depict persons in familiar situations.

Each picture was then examined on the basis of six criteria culled from the literature and adapted to our purposes. Thirty-two pictures were discarded after critical scrutiny based on the following criteria.

1. *The picture situation should relate to the particular problem under investigation.* This not only was the particular theme of the present research but has some support from Vernon's study of adolescent fantasies [10], in which he concludes, "It appears that the response was to a considerable extent a function of the picture shown."

2. *The pictures should represent situations familiar to the subjects.* Symonds [9] claims that pictures of natural, everyday, and present-day settings yield better stories than fanciful pictures. See also Haggard [1] and Harrison [2].

3. *The pictures should contain a central figure of the same sex and approximately the same age as the subjects.* Murray [3] found that "in most of the pictures there should be at least one person with whom the subject could easily empathize and identify himself." See also Rotter [8] and Symonds [9].

4. *The pictures should not contain characters or situations which are intrinsically humorous.* Humor was guarded against, perhaps needlessly, as cartoon pictures were used with captions deleted. Since humor is *de gustibus,* no claim can be made that it was eliminated, nor did we find that our own notions of humor made much difference in the character of obtained responses.

5. *The pictures should contain a minimum of detail.* Symonds [9] found this criterion to correlate most highly with the "goodness" of his pictures. Minimum detail is also a means of obtaining ambiguity. See Rotter [8] for further amplification.

6. *The pictures should depict an event episodic in nature but should not tell a complete story.* See Murray [3] and Symonds [9].

The 18 pictures remaining were then presented as a trial run to 12 subjects approximating the characteristics of the group selected for further study. As was suspected, the obtained results indicated that those pictures not including *both* child and parent figures often failed to evoke stories with child-parent content. Three such pictures were eliminated, but one was retained in the final series of 15 (a child praying by his bed) since it had evoked a large number of suitable responses in the trial run.

There is little doubt that critical inspection based on the now fairly well accepted criteria for "goodness" of projective pictures will yield a *group* of pictures which are better than other pictures rejected by the inspection technique. However, it is unlikely that the available criteria are sufficiently precise to indicate which of a fairly large group of pictures are *most likely* to evoke meaningful stories; nor can inspection techniques provide an acceptable ranking of a given group of "good" pictures. Although the question of goodness is in reality inextricably bound to the basic problem of validity, no known technique for ascertaining validity of imaginative productions is known to the writers, as such validity must reside in the realm of the imaginative and subjective and cannot be determined by behavioral observations. A pragmatic validity which pertains to the "usefulness" of imaginative productions *is,* however, behavior-linked, and probably can be determined by the usual methods for determining technical usefulness even though the resulting quantitative statements are admittedly approximate and open to various sources of error. Neither of these basic validities has been touched upon in the present study, since the history of projective techniques indicates that for the time being, at least, practical results can be obtained from unproved methods in the hands of experienced clinicians.

Thus, the clinician was adopted as the investigative tool, and protocols from the 15 selected pictures were submitted to the judgment of five relatively experienced judges (graduate students who had used the TAT in clinical internships and practica) who had been trained in the application of two rating scales for judging the "response usefulness" of

individual stories in terms of the avowed purpose of investigating child-parent relationships. The stories submitted to the judges were obtained from 30 boys between the ages of 10-0 and 12-9 institutionalized at the Ohio Bureau of Juvenile Research for observation and diagnosis because of a wide variety of behavior problems. The mean IQ on individual intelligence tests was 94, with a range from 62 to 119.

Protocols were obtained from the children by the senior author after intensive training in administration of the pictures to the trial group. Instructions to the subjects were a modification of those advocated for children by Rotter [8] and were essentially as follows:

> I have a set of pictures here, and I want you to make up a story about each one. You may make the story as long or as short as you wish, just so you tell me a story. You might tell me who the people in the picture are; what they are talking or thinking about; what they have been doing or are going to do. While you are telling the story, I am going to write it down. If you go too fast, I may tell you to slow down a bit.

Since electrical recording techniques were not feasible, almost verbatim recording of all responses was obtained by abbreviation and elimination of side comments. In order to minimize the effect of serial position, the pictures were presented in reverse order to every second child.

The 450 stories resulting from 30 subjects responding to the 15 stimulus pictures were independently rated by five judges against each of two rating scales. Mimeographed forms and predetermined order of judgment were used in order that the ratings for one subject or for one picture would not substantially influence the ratings for the next picture or the next subject.

The variables essential for the rating of the "usefulness" of story protocols are unknown, but several can be assumed from logical analysis of the purpose of projective pictures as interview techniques. After considerable weighing of alternatives, two variables were finally selected as meeting the following criteria: (1) It must be a variable which will be suitable for the immediate purpose of the investigation. (2) It must appear to have a low correlation with other variables employed. How many variables to use and what criteria are most important in their selection are, of course, matters for considerable empirical research which can yield exact results only when full knowledge of the other variables involved in the clinical use of imaginative productions are more clearly defined. In the absence of such data, logical hypotheses seem appropriate at least as preliminary tools.

The two rating scales selected were tentatively entitled "Emotional Involvement" and "Conceptual Relationship." Detailed illustration of only the latter is presented below since analysis revealed that the two scales, contrary to expectations, were rather highly correlated (*rho* was .79). The instructions given to the judges described the five-point scale for Emotional Involvement as follows.

> It is believed that the degree of emotional involvement or feeling tone invested in a particular story gives some indication of the amount of projection and empathy invoked by that stimulus picture. Since this is an investigation of child-parent relationships, the emotional involvement rated must be that revolving about a child-parent constellation. In addition to plots of an obvious emotional nature, involvement should be looked for in those productions that are quite original, where there appear sudden gaps, blocks, contradictions, and inappropriate wish-fulfilling endings. It is important to consider "positive," pleasant emotional material as well as negatively toned material.

The following section is a slightly abbreviated version of the instructions given to the judges for the second scale. The response examples are not those employed originally but are actual responses with ratings determined by the judges of the present study.

CONCEPTUAL RELATIONSHIP SCALE

> It is believed that, in addition to Emotional Involvement, at least one other factor patently plays an important role in determining picture response usefulness. This factor, "Conceptual Relationship," when used in connection with the particular subject under investigation (child-parent relationship), translates into the following question:
>
> "How significant do you believe each response to be in understanding the child's attitudes toward his parents and how he conceives of their attitudes toward him?"
>
> Your task is to rate, against the following five-point scale, the amount of significance each of the fifteen stories given by a particular child has in

understanding the child's conceptual relationship with his parents.

Although emotional involvement will frequently participate in the consideration of a story's significance, emphasis upon this factor should be avoided. Often, stories containing little or no emotionality may be quite significant in revealing child-parent relationships. In the interest of inter-judge consistency it is also requested that you avoid the use of "depth" interpretation and interpret the stories literally. Finally, it is realized that it may be very difficult to point to the significance, or lack of it, in particular stories since there are frequently interpretive interrelations between stories. Nonetheless, you are asked to use your best judgment.

1. *First Level.* Probably of no use in understanding the child-parent conceptual relationship. Example:

   (Picture 8) The little boy is saying his prayers. He is on his knees beside his bed. He finishes his prayer. He was praying about Christmas. What he was going to get.

2. *Second Level.* Might be of some small value in understanding the child-parent conceptual relationship. Example:

   (Picture 11) He says, "Let me see the globe in the other room." The father says, "All right. Come with me, you pain-in-the-neck." He says, "What's this state?" The father says, "It's Cleveland, Ohio." The boy says, "What is this city?" He says, "That's New York." The father says, "This is the biggest ocean in the world. The Pacific." The boy says, "Yeah?" The father says, "I bet you don't know what the widest river is." The boy says, "The Amazon. Right?"

3. *Third Level.* Probably of medium usefulness in understanding the child-parent conceptual relationship. Example:

   (Picture 13) The mother wants her son to go in there and ask the father to take him to the show but he is scared to ask him. The mother says, "He won't bite you. All he can do is say 'No'." So he goes in there and asks his father to take him to the show. The father says, "No." So he says, "If you don't take me to the show I will tell mother how late you stayed out last night."

4. *Fourth Level.* Probably of above-average usefulness in understanding the child-parent conceptual relationship.

   (Picture 4) He broke a pitcher and ran and hid. His father got mad at him and chased him and never found him. He called him and the kid ran out the door and ran away. About four hours later he came back and his dad and mother weren't at home. He stole five hundred dollars. When his dad saw it was gone he called five hundred cops. They went for a hunt. They finally found him and he was dead—with a bullet. He was eighteen years old.

5. *Fifth Level.* Probably exceedingly useful in understanding the child-parent conceptual relationship.

   (Picture 7) The mother said, "Don't get too close to the fire 'cause your hair will catch on fire and your hair will electrocute you." The boy said, 'Why should it? I'll go stick my head in the fire and show you." He stuck his head in the fire and his mother tried to save him and she fell in the fire. Then his father fell in the fire, too. They all got burned up.

## RESULTS

The reliability of the judges in rating the pictures was a matter of primary concern. This problem was approached in two ways. The first was a simple rank-difference correlation based on inter-judge picture ranks where picture rank was obtained from a raw score based on simple summation of levels as rated by the judge on each rating scale. The median *rho* obtained was .84, with a surprisingly small range from .78 to .93. Nonetheless, these findings revealed differences between judges' means for each picture, which may or may not have been attributable to chance. An analysis of variance was therefore performed on the picture response ratings which showed that the picture, judge, and scale means differed significantly. Also, there appeared statistically significant variations among the interactions of these three factors. Despite the fact that the means of the judges varied significantly, this appeared to be primarily a limiting factor, since it was further shown that a significant $F$ ratio existed between the picture variance and the judge with picture interaction. Thus, the difference between picture means could not be accounted for by the variance in judges' means. As is commonly found in ratings based on judgments, the agreement between judges was found to be high, although statistically significant differences between judges were manifest. These did not, however, account for the statistically significant differences between picture ranks. The differences between judges can probably be further reduced by more intensive training. On Scale 2, Conceptual

Relationship, the conversion of raw ratings to ratings based on standard scores might further reduce the inter-judge variability.

By summing the scores given a picture by all five judges on one scale, a "picture mean" was obtained for each picture on each scale. Since there were only 15 items to be correlated, a rank difference correlation was used to measure the relationship between the scales despite the fact that the analysis of variance had shown that the two scales actually measured different aspects of picture usefulness. The resulting *rho* of .79 shows that although the two rating scales are probably measuring somewhat different variables, the variance is small and the aspects of picture usefulness measured are closely related.

The analysis of variance showed that the differences between picture means (based on summated ratings of all four judges using *both* rating scales) could not be accounted for by the unreliability of the judges. Thus, it appears clear that the pictures themselves must have had a major influence on the different scores obtained, confirming the hypothesis that in a given series of similar pictures some will, generally, prove more useful than others.

TABLE 1

STATISTICAL DIFFERENCES BETWEEN PICTURE MEANS (COMBINED SCALES)

| Rank | Picture No. | Mean | *t*-Ratio* |
|---|---|---|---|
| 1 | 4 | 3.51 | |
| | | | 2.97 |
| 2 | 13 | 3.19 | |
| | | | 0.19 |
| 3 | 3 | 3.17 | |
| | | | 1.39 |
| 4 | 12 | 3.02 | |
| | | | 0.92 |
| 5 | 7 | 3.01 | |
| | | | 0.00 |
| 6 | 5 | 3.01 | |
| | | | 1.11 |
| 7 | 1 | 2.89 | |
| | | | 0.46 |
| 8 | 14 | 2.84 | |
| | | | 0.56 |
| 9 | 9 | 2.78 | |
| | | | 1.02 |
| 10 | 6 | 2.67 | |
| | | | 0.74 |
| 11 | 15 | 2.59 | |
| | | | 0.37 |
| 12 | 10 | 2.55 | |
| | | | 1.02 |
| 13 | 2 | 2.44 | |
| | | | 2.69 |
| 14 | 8 | 2.15 | |
| | | | 3.53 |
| 15 | 11 | 1.77 | |

*A t of 2.61 is significant at the .01 level, while a t of 1.98 is significant at the .05 level with df = 1.49.

By arranging the pictures in rank order in terms of their total mean score, *t*-tests of significance of differences permit evaluation of the comparative usefulness of each picture. Table 1 shows a clearly significant difference between the mean scores of picture 4 and picture 13, the two highest-ranking pictures. There is no significant difference between pictures 13 and 3, which rank second and third, respectively. However, there appear to be reliable differences in usefulness between the first eight pictures and the last seven pictures in this particular sample, since there is a highly reliable difference ($t = 6.22$) between the picture ranked as 1 and the picture ranked as 9. Pictures 8 and 11, ranking 14 and 15, are clearly the least useful statistically. Examination of these two shows that they probably fail to meet adequately the criteria for selection as evokers of stories about child-parent relationships. The statistical findings are thus partially validated in so far as questionable pictures are concerned.

One clear-cut result of this study points to a more general conclusion. No picture which barely meets the subjective criteria was found to be always low in usefulness, nor, conversely, does even the best picture always act as a stimulus to "good" stories. For example, picture 11, ranking 15, was rated as producing better-than-average stories about 5 per cent of the time, while picture 4, ranking first, was rated as producing poorer-than-average stories about 15 per cent of the time.

Further analyses failed to reveal any significant correlation between age and response level, or intelligence and response level. With the highly curtailed age range in our population, the absence of correlation is not unexpected. Another finding confirmed a tentative hypothesis that length of response bears some relation to rated usefulness, as a *rho* of .83 was obtained between the number of lines per response and the rated usefulness of that

response. Since, however, the recording of stories was not fully controlled, and as there was some question as to the complete sophistication of judges in the rating technique, the hypothesis is still tentative, although it is one which is unlikely to be grossly inaccurate. If further research upholds the hypothesis at the level of the obtained *rho,* a very rapid procedure for selecting "good" pictures could be evolved using only Rating Scale II (Conceptual Relationship) and an accurate measure of length of story elicited under standard conditions.

## DISCUSSION

The present study was undertaken as a methodological attempt to demonstrate the possibility of developing a set of projective pictures which would provide tools for obtaining clinically useful information on a given topic. There is no reason to believe that the best eight cartoon pictures of the present study are those which would generally prove to be most useful in obtaining information on child-parent relationships from all subjects of the same age range. Nonetheless, it is believed that the method of obtaining a fairly large group of pictures which meet a set of generally accepted subjective criteria, followed by a systematic rating of obtained stories by qualified judges, provides a fairly rapid means of obtaining numerous sets of projective pictures which will prove *generally* useful in most clinical situations. Although we speak of this method as being "rapid," it is by no means a task to be completed overnight. Time could probably be saved by using fewer than five judges and by using one, rather than two, rating scales since the two scales of the present study are rather highly correlated. Even after such short cuts, however, the problem of obtaining empirical information on the usefulness of the pictures seems essential and does involve considerable time.

No special faith is placed in the statistical methods applied to the ranking of pictures in the present study. Other statistical procedures could easily be adapted but would probably provide quite similar results. If special faith is required, it probably rests on the notion that sets of projective pictures empirically tested for eliciting stories on particular topics can be developed in a fairly short time. Such sets of pictures adapted to particular age and sex groups would probably be of great clinical advantage as "interview-aids" or as tools to plumb more thoroughly the depths which are only partially exposed by the TAT survey techniques. It may be some years in the future, but it is not too unreal to expect that someone will publish a large-sized "drawer" of pictures, indexed as to topic, age, and sex. Such a "drawer" would be a welcome supplement to the clinician's other equipment and should provide him with much more flexible tools with which he can hope to reduce the errors and omissions implicit in clinical analysis.

Topics which are general enough in their interest to clinicians might well include all of the following sets of pictures available for each age and sex group—child-peer relations, sibling relationships, heterosexual relationships, school relationships, job or occupational relationships, parent-child relationships with pictures particularly selected for adults, husband-wife relationships, employer-employee relationships, etc.

## SUMMARY

In diagnostic interviewing, projective pictures designed to elicit stories on a specified topic can prove useful to supplement general survey techniques such as the TAT. In order to explore the methodology, 15 pictures from an original selection of 50 thought to have bearing on child-parent relationships, were administered to 30 institutionalized boys. The stories obtained were rated by five judges on two rating scales entitled "Emotional Involvement" and "Conceptual Relationship." Statistical analysis revealed that the 15 pictures could be ranked in order of usefulness, with statistically significant differences in rated usefulness appearing between the eight top-ranking pictures and the lowest seven. It is suggested that similar sets of pictures, empirically and statistically tested, would be of use to clinicians in investigating specific topics which now can be handled only by ordinary interview technics.

*Received January 27, 1951.*

## References

1. Haggard, E. A. A projective technique using comic strip characters. *Character & Pers.*, 1942, **10**, 289–295.
2. Harrison, R. Studies in the use and validity of the Thematic Apperception Test with mentally disordered patients. *Character & Pers.*, 1941, **9**, 122–138.
3. Murray, H. A., *et al. Explorations in personality.* New York: Oxford Univ. Press, 1938.
4. Rapaport, D. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1946. Vol. II.
5. Rosenzweig, S., and Sarason, S. Fantasy in personality and its study in test procedures. *J. abnorm. soc. Psychol.*, 1942, **37**, 40–54.
6. Rosenzweig, S. The picture association method and its application in a study of reactions to frustration. *J. Personality*, 1945, **14**, 3–23.
7. Rosenzweig, S., Fleming, Edith E., and Rosenzweig, Louise. The children's form of the Rosenzweig Picture-Frustration Study. *J. Psychol.*, 1948, **26**, 141–191.
8. Rotter, J. B. Thematic apperception tests: Suggestions for administration and interpretation. *J. Personality*, 1946, **15**, 70–92.
9. Symonds, P. M. Criteria for the selection of pictures for the investigation of adolescent phantasies. *J. abnorm. soc. Psychol.*, 1939, **34**, 271–274.
10. Vernon, M. D. The relation of cognition and phantasy in children. *Brit. J. Psychol.*, 1940, **30**, 273–295; **31**, 1–22.

# SOME EFFECTS OF COLOR ON THEMATIC FANTASY[1]

GLEN A. BRACKBILL[2]
UNIVERSITY OF CALIFORNIA, LOS ANGELES, AND
VETERANS ADMINISTRATION HOSPITAL, LOS ANGELES

THE diagnostic and dynamic implications of individual reactions to color in projective test stimuli is a topic which has recently received considerable attention in the literature of diagnostic testing. Most experimental studies in this area have been based on analyses of responses to the colored cards of the Rorschach ink blot technique [1, 8, 11, 12, 14, 18]. This is to be expected inasmuch as the Rorschach method involves the systematic use of reactions to color as an integral part of the test analysis.

Reports on the validity of interpretations of reactions to color in the Rorschach test are conflicting. Although further study is needed before any definite conclusions can be reached, a number of researches seem to indicate that reactions to color are, at least in part, a function of personality [7, 10, 15, 18]. The writer suggests that further information pertinent to this topic could be gained by investigating some of the effects of using color in fantasy-evoking projective techniques such as the Thematic Apperception Test (TAT) [9] and the Make-a-Picture-Story-Test (MAPS) [16]. This study was undertaken in an attempt to discover some of these effects. It was hypothesized that if color does act as an emotional stimulant, the responses of neurotic persons to color in the TAT will be significantly different from reactions of normal control subjects. If such differences are found, the use of color in fantasy stimulus material would add to the test another dimension for personality evaluation.

The question of the use of color in the TAT is particularly interesting in view of a recent study reported by Garfield and Eron [6]. They found that the mood or emotional tone of TAT stories told by 25 college students and 25 schizophrenic patients was more a function of the stimulus properties of the test pictures than of the personality of the subject. Undoubtedly, most clinicians working with the TAT or other similar devices have heard numerous comments from their subjects concerning the depressiveness of the gray and black backgrounds of the pictures. The dysphoric or darkness characteristics of the pictures probably impose a limitation on the subject's freedom to express his own feelings in the stories he tells. Adding color or tint to the pictures should free the subject from such limits. Projection of his own prevailing mood should be easier, then, for two reasons: (1) the emotion-stimulating quality of the color itself and (2) freedom from the limits imposed by the lifeless black and gray of the pictures. Thus, with colored pictures we would expect to find the emotional tone of the story less a function of the card and more a reflection of the subject's own mood.

## METHOD AND PROCEDURE

The experimental group was made up of 50 subjects from two sources. Twenty-eight were patients at a mental hygiene clinic. In most cases, the patients were under treatment by psychologists or psychiatric social workers. The criterion for inclusion in this group was a diagnosis of psychoneurosis or some form of character disorder. Diagnosis was made by the therapists and was based primarily upon impressions gained in therapeutic interviews. In some cases the results of a battery of pyschological tests were considered in the final

[1]Published with permission of the Chief Medical Director, Department of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed or the conclusions drawn by the author. Based on a dissertation submitted to the University of California, Los Angeles, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer gratefully acknowledges his indebtedness to the committee who supervised his research.

[2]Now at the University of Colorado.

diagnosis. The remaining 22 subjects were patients at a Veterans Administration neuropsychiatric hospital. They were included in the experimental group if they had an admission diagnosis of character disorder or one of the pyschoneuroses. The diagnosis was checked when the patient was discharged, and any behavior of the patient during his hospitalization that suggested the admission diagnosis was in error led to a re-examination of the case. If there were indications of psychosis, the patient was dropped from the group and another was added.

It is generally known that differential diagnoses are quite unreliable as criteria for group homogeneity. However, in general, the distinction between psychosis and psychoneurosis is more easily established. Dorcus and Shaffer point out:

> It is entirely conceivable, therefore, that what is thought to be abnormal by one investigator may be considered quite within the range of normal by another. There is, however, considerable difference between the well-developed psychoneuroses and the normal and between full-fledged psychoses and the neuroses [3, p. 354].

Differential diagnoses included in the experimental group are shown in Table 1. Other

TABLE 1

DISTRIBUTION OF DIAGNOSTIC CATEGORIES

| Diagnostic Classification | Number of Subjects E1 | E2 | Total |
|---|---|---|---|
| Anxiety Reaction | 10 | 11 | 21 |
| Anxiety Hysteria | 5 | 4 | 9 |
| Obsessive Compulsive | 5 | 2 | 7 |
| Immaturity Reaction (Passive Dependence) | 3 | 5 | 8 |
| Somatic Reaction | 2 | 2 | 4 |
| Mixed Neuroses | 0 | 1 | 1 |
| Total | 25 | 25 | 50 |

TABLE 2

COMPARISON OF CHRONOLOGICAL AGE FOR EXPERIMENTAL AND CONTROL GROUP

| | Experimental Group ($N$=50) | Control Group ($N$=50) |
|---|---|---|
| Mean | 33.0 | 30.5 |
| Sigma | 7.9 | 8.1 |
| Range | 21–54 | 21–51 |

TABLE 3

COMPARISON OF IQ FOR EXPERIMENTAL AND CONTROL GROUPS

| | Experimental Group Wechsler-Bellevue ($N$=24) | Control Group Otis IQ ($N$=50) |
|---|---|---|
| Mean | 110.6 | 112.5 |
| Sigma | 15.7 | 10.4 |
| Range | 75–135 | 84–145 |

data about this group are presented in Tables 2 and 3.

The control group also came from two main sources. Twenty-six were volunteers from a Veterans Administration general medicine and surgery hospital; the remaining subjects were recruited from a young peoples' social organization. Nonhospitalized controls were obtained in order to hold constant the factor of hospitalization. The clinic patients in the experimental group, of course, lived at home. The hospitalized control subjects were patients on the EENT service. Most of them were pretonsillectomy cases that had been in the hospital from three to five days. No patients who suffered what appeared to be chronic disturbances of EENT functions were used. In order to introduce some objective criterion of normalcy, the Cornell Selectee Index, Form N [20], was used, with an arbitrary cutoff point of 12 items answered in the aberrant direction and no "stop" items. Subjects who suffered any kind of permanent physical disability such as facial disfigurement, loss of limb, or paralysis were excluded.

The two groups were comparable in terms of (1) age, (2) intelligence, (3) educational level, (4) veteran status, and (5) hospitalization. They differed in terms of (1) diagnosis and (2) number of aberrant responses to the Cornell Selectee Index.

The Cornell Selectee Index was administered to all control subjects prior to any other tests. Visual inspection of the record after the subject had completed the test was sufficient to determine whether or not further testing was in order. This was possible since a "yes" response to all but five items is an aberrant response. The scoring was later checked by the use of a hand stencil. Twenty-five of the experimental subjects also filled out the Index. Intelligence tests were administered to all

control subjects and to as many of the experimental subjects as the exigencies of clinical research permitted (see Table 3).

The TAT (1943 Edition) was used to study the influence of color on fantasy test productions. Cards 1, 2, 3BM, 4, 6BM, 7BM, 8BM, 11, 13MF, 17BM, 18GF, and 16[3] were selected and administered in the order named. A set of cards was given to a commercial artist who tinted them in oils. Tints were selected in view of the findings of Washburn [19] and Dorcus [2] that tints are more effective emotional stimulants than are bright or saturated colors. Every effort was made by the artist to avoid changing the facial expressions of any of the figures or defining the outlines of the vague objects in some of the pictures. Flesh tones were used on all human figures, and backgrounds were tinted with colors thought appropriate by the writer and the artist.

Two sets of cards were used in the study. In set 1 the first six cards were colored, and the second six were the original black and white. In set 2 the first six cards were the original black and white, and the second six were colored. All tests were administered individually by the writer. In all, the accumulated data represented 1,200 stories. The same test instructions were used for all subjects and were those suggested by Rotter [13, pp. 70-72]. Sets 1 and 2 of the modified TAT were administered in an ABBA manner to both experimental and control groups. However, analysis of the results indicated that order of presentation was not a significant factor contributing to the variables measured in this study. The 100 subjects were then divided into four groups of 25,[4] as shown in Table 4.

Immediately following the TAT, a questionnaire was filled out by the experimenter who asked the questions orally and recorded the responses. Part I of the questionnaire contained items of personal data such as age, education, and vocation. Part II called for the subject's preference for the colored or uncolored cards and an inquiry into the possible sources for the stories told, i. e., personal experience, books read, movies, etc.

TABLE 4
COMPOSITION OF SUBGROUPS

| Administered | Control Group | *N* | Experimental Group | *N* | Total |
|---|---|---|---|---|---|
| Set 1 Colored | C1 | 25 | E1 | 25 | 50 |
| Set 2 Colored | C2 | 25 | E2 | 25 | 50 |

## RESULTS

Since the discovery of possible effects of color on the mood of TAT stories was of primary interest in this study, it was necessary to establish some means of judging what mood was predominant in each of the stories. In order to establish some measure of reliability of the judgments, the writer enlisted the aid of five graduate students who had had some experience in the clinical use of the TAT. Each judge classified 48 stories picked by the use of random numbers, from the 1,200 stories, so that 24 were from the experimental group and 24 were from the control group. Each judge was asked to sort the stories into four possible categories, using the following descriptions adapted from a list suggested by Schorr [17]:

*Depressive*

1. The story is sad toned.
2. There is marked conflict among the characters or between the characters and the environment.
3. There is implied helplessness, hopelessness or gloominess.
4. Main characters are depressed most of the time, act under coercion from within self or without, are anxious.

*Destructive*

1. The story contains violence, murder, suicide, rape, crime.
2. There are themes of world catastrophe, i.e., fire, flood, end of the world.
3. Story is ended with death of the hero.

*Intellectualized*

1. Story is highly factual discussion.

[3]The selection of cards used was based, in part, on suggestions made by Dr. Bruno Klopfer.

[4]For purposes of statistical analysis the stories told to the second six cards in each subgroup of both experimentals and controls were interchanged, i.e., the stories told to the uncolored cards of group E1 were placed in group E2, and stories to the six colored cards in group E2 were placed in group E1. The same procedure was followed in the control group. This resulted in experimental and control groups which could be dealt with as if the subjects had responded to all colored or uncolored cards.

TABLE 5

SIGNIFICANCE OF DIFFERENCES IN MOODS

| Group | Mood | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Depressive | | | | Intellectual | | | | Destructive | | | | Other | | | |
| | % | Diff | CR | P | % | Diff | CR | P | % | Diff | CR | P | % | Diff | CR | P |
| E1 | 64 | | | | 14 | | | | 3 | | | | 17 | | | |
| | | 10 | 1.78 | .06 | | −2 | — | — | | −1 | — | — | | −9 | 1.98 | .04 |
| E2 | 54 | | | | 16 | | | | 4 | | | | 26 | | | |
| C1 | 52 | | | | 7 | | | | 5 | | | | 36 | | | |
| | | 3 | 0.53 | .60 | | −8 | 2.28 | .02 | | −6 | 2.50 | .01 | | 11 | 2.04 | .04 |
| C2 | 49 | | | | 15 | | | | 11 | | | | 25 | | | |
| E1 | 64 | | | | 14 | | | | 3 | | | | 17 | | | |
| | | 12 | 2.14 | .02 | | 7 | 2.06 | .04 | | −2 | — | — | | −16 | 3.13 | .01 |
| C1 | 52 | | | | 7 | | | | 5 | | | | 36 | | | |
| E2 | 54 | | | | 16 | | | | 4 | | | | 26 | | | |
| | | 5 | 0.87 | .30 | | −1 | — | — | | −7 | 3.18 | .002 | | 1 | — | — |
| C2 | 49 | | | | 15 | | | | 11 | | | | 25 | | | |

2. There is little or no reference to interaction among the characters, no evidence of "emotional involvement."
3. The story is mostly a description of the picture.

*Other*

1. Moods or feelings which cannot be placed in one of the above three categories are to be classified as Other.

To check the agreement of the writer with the other five judges, the cosine tetrachoric correlation suggested by Dunlap and Kurtz [4] was used. Fifteen correlations were computed involving comparison of the experimenter's scoring with that of each of the five judges on the three moods, Depressive, Destructive, and Intellectual. The correlations ranged from .60 to .94. The category Intellectual showed the lowest amount of agreement. The mean *r* was .80 for the 15 comparisons. These *r*'s compare favorably with those found in the literature. Garfield and Eron [6] reported tetrachoric *r*'s varying from .75 to .92 among judgment of mood, outcome, and activity-passivity of characters.

The mood analysis was made on the basis of the first six stories told by each of the four subgroups. Results of this analysis are presented in Table 5. The differences most pertinent to the present study are those between E1 and E2, and, correspondingly, between C1 and C2 (see Table 5). Use of the colored TAT resulted in a greater number of Depressive stories among the neurotic subjects than the use of uncolored cards. Increase in Depressive stories was accompanied by a significant decrease in the percentage of stories classed as Other. The control group did not show an increase in Depressive stories, but there were significant decreases in the percentage of Intellectual and Destructive stories, with an increase of stories classed as Other. Thus, it might be stated that the neurotic subjects became more depressed in the face of the emotional stimulation provided by the color. The control or normal group, however, reacted with less disturbing feeling, as evidenced by a shift to moods classified as Other. The category, Other, contained happy, neutral, or shifting moods.

Two other measures which demonstrate some effects of color on the emotional tone of TAT stories were used. They were reaction time,[5] and productivity as measured by number of words per story. The analysis of variance technique was used, with the cards constituting the columns and the various sub-

[5]For purposes of this study, reaction time was defined as the interval from the time the subject took the card to the beginning of his first comment.

groups the rows. Stories to all 12 cards were included. The analysis indicated that the longer mean reaction time to the colored cards than to the uncolored cards was significant (CR = 8.53, $P = .001$) for the neurotic group. However, when the comparison was made for the control subgroups, the difference was not significant. Taken as a whole, the neurotic group showed significantly longer reaction times than did the normal group ($P = .001$.).

The analysis, using productivity as one of the variables, did not reveal such marked differences. The significance of the difference between E1 and E2 was reliable at the .17 level of confidence. There was, again, no significant difference between the control groups. However, it is interesting to note that when the experimental and control groups as a whole were compared, the neurotic subjects were significantly more productive than the normal subjects ($P = .001$).

The introspective reports of preference for colored cards over noncolored cards were quite different for the two major groups. Seventy-four per cent of the normal group reported a preference for the colored cards, 16 per cent for uncolored, and 10 per cent felt that it made no difference. However, only 66 per cent of the neurotic subjects preferred colored cards, 30 per cent chose uncolored, and 4 per cent had no preference. The difference in preference for colored cards is not significant, but the difference of 14 per cent in the choice of uncolored cards approaches significance ($P = .08$).

## DISCUSSION

The use of colored stimulus cards in the TAT apparently has several interesting effects on the responses produced. Furthermore, these effects vary with personality adjustment. The increase in the number of depressed stories might be explained in terms suggested by Wallen [18], who felt that color in the Rorschach test may stimulate disturbing affective associations. The colored TAT cards would perhaps offer an even stronger stimulation to affective associations since the color appears in connection with a psychosocial situation depicted in the cards. Certainly, the subjects who were members of the experimental group manifested many symptoms of anxiety and depression stemming from difficulties in both personal and social adjustment. The following excerpts from admission interviews would support this statement.

*Case 2.* Patient complains that he is afraid of something but he doesn't know what it is. He is worried about low grades in school, fears failure.

*Case 7.* Patient states he has no goal in life, doesn't know what he wants to do. Expressed feelings of guilt and depression over not being sent overseas.

*Case 14.* Complains of being tired, tense and worried about his inability to gain weight. Can't concentrate on his work, can't sleep or relax.

*Case 18.* Concern about symptoms of heart racing, dizziness, constant anxiety, afraid of something, but not sure what.

*Case 52.* Patient has been increasingly nervous and depressed. Has been unable to hold a job. Made a suicidal gesture. Has numerous fears.

These subjects can be described clinically as depressed and anxious. Thus, it would be reasonable to expect that stimulation which produces an expression of the subject's prevailing mood would reveal this depressive affect. The results of the investigation, as shown in the greater number of Depressive stories told to colored cards, supports the hypothesis that the mood of stories told to colored cards is a more accurate index of the subject's affective tone than are moods of stories told to black and white cards. Examination of Table 5 will show that the experimental group could not be distinguished from the control group in terms of mood based on stories told to black and white cards but could be distinguished using stories told to colored cards. The same situation held in the consideration of reaction time and productivity.

It is interesting to note some of the comments made by the neurotic subjects when they were asked to explain their preference for uncolored cards. Such comments as the following were given. "The color seems dirty." "It is bad art." "It is easier to see the expression in black and white." "The color is misleading." "The color seems made-up." These comments indicate discomfort or disturbing feeling aroused by the color and can be taken as further evidence for the

hypothesis that color stimulated the predominantly depressive mood of the neurotic subjects which was then expressed in the story told. It would, however, be necessary to study the effects of color in each card in order to be able to decide whether colored cards would meet the criticism implied in Eron's conclusions [5].

The control group also reacted to color with a change of mood. However, the shift took the form of a significant reduction in Destructive stories and a rise in the percentage of stories classed as Other. It is a common fact that most "normal" people associate darkness and gray with tragedy, destruction, and gloom. Witness the markedly effective use of black and white in Olivier's motion picture production of *Hamlet*. Many agree that this usage intensified light and shadow, and served admirably to set the stage for Shakespeare's great tragedy. Use of colored cards seemed to have the effect of arousing less of the destructive and tragic association and more of what might be termed lighter or happier associations, as reflected by an increase in the category Other.

The longer reaction time to colored cards can be interpreted to support the hypothesis that disturbing affect has been aroused. Significant increases in reaction times in word association and other clinical tests has long been recognized as an indicator of emotional blocks or disturbed feelings. The fact that the normal subjects did not show a longer reaction time to colored cards may also be interpreted to support the thesis that the affect aroused by color among normals was not as disturbing as that produced in the neurotic subjects.

While productivity was significantly greater among the experimental subjects than among the control group, the effects of color on this variable were not as marked. For the neurotic group, $P = .17$. Several hypotheses could be suggested regarding the difference between the major groups. We might have expected that, being less anxious, the normal subjects would be more spontaneous in fantasy production with consequent longer stories. However, it might also be expected that the emotionally disturbed group told longer stories because lack of self-confidence, and anxiety about the adequacy of their performances led them to elaborate and justify their statements about the pictures, thus producing longer stories. Another possibility is that the length of stories told by neurotics is the result of a tendency of some such persons to deal with the test situation by violating the instructions and engaging in a prolonged description of the picture. That the color has any relation to this variable is questionable.

Obviously, in an exploratory investigation such as this, more questions are raised than answered. The hypotheses suggested here are based on the findings and have not been tested by the experiment. However, it seems rather clear that there are interesting possibilities in the further investigation of relationships between color and emotionality and their clinical diagnostic use.

## SUMMARY

This study was undertaken to investigate possible effects of using colored stimulus material in such fantasy tests as the MAPS and TAT. Fifty subjects diagnosed as psychoneurotic and 50 normal control subjects were studied. One half of each group was given six colored TAT cards followed by six uncolored cards. For the remaining two groups the order of presentation was reversed. The following results were noted: (1) Subjects diagnosed as psychoneurotic told more depressed stories to chromatic than to achromatic cards. (2) Normal subjects showed no change in the number of Depressive stories but did have a significant decrease in Intellectual and Destructive stories to chromatic cards. This resulted in an increase of stories classed as Other. (3) The neurotic subjects had a significantly longer mean reaction time to chromatic cards than to achromatic cards. (4) The control groups showed no difference in reaction time. (5) Productivity as measured by number of words per story showed a tendency to increase with the use of color in the neurotic group. In general, the patient group was significantly more productive than the normal control group. It was suggested that the findings could be ascribed to the effectiveness of color in stimulating the expression of the subject's own prevailing mood in his stories. Color may

play the part of arousing unpleasant emotional associations. It was further suggested that the use of color might, in part, meet the criticism that the mood of TAT stories is more a function of the test card than it is of the personality of the subject.

*Received January 2, 1951.*

REFERENCES


1. ALLEN, R. M. A simple method of validating color and shading shock. *J. consult. Psychol.*, 1948, **12**, 360.
2. DORCUS, R. M. Color preference and color associations. *Ped. Sem.*, 1926, **33**, 399–434.
3. DORCUS, R. M., AND SHAFFER, G. W. *Textbook of abnormal psychology.* Baltimore: Williams & Wilkins, 1945.
4. DUNLAP, J., AND KURTZ, A. *Handbook of statistical nomographs, tables and formulas.* Yonkers, N. Y.: World Book Co., 1932.
5. ERON, L. D. Frequencies of themes and identification in the stories of schizophrenic patients and non-hospitalized college students. *J. consult. Psychol.*, 1948, **12**, 387–395.
6. GARFIELD, S. L., AND ERON, L. D. Interpreting mood and activity in TAT stories. *J. abnorm. soc. Psychol.*, 1948, **43**, 338–345.
7. GOLDSTEIN, K. Some experimental observations concerning the influence of colors on the function of the organism. *Occup. Ther. Rehabilit.*, 1942, **21**, 147–151.
8. LAZARUS, R. The influence of color on the protocol of the Rorschach test. *J. abnorm. soc. Psychol.*, 1949, **44**, 506–516.
9. MURRAY, H. *Thematic Apperception Test manual.* Cambridge: Harvard Univ. Press, 1943.
10. OESER, O. A. Some experiments on the abstraction of form and colour. *Brit. J. Psychol.*, 1932, **22**, 200–215.
11. OESER, O. A. Some experiments on the abstraction of form and colour: Part II. Rorschach tests. *Brit. J. Psychol.*, 1932, **22**, 287–323.
12. RORSCHACH, H. *Psychodiagnostics.* Berne: Hans Huber, 1942.
13. ROTTER, J. B. Studies in the use and validity of the Thematic Apperception Test. I. Method of analysis and clinical problems. *Character & Pers.*, 1940, **9**, 18–34.
14. SAPPENFIELD, B. R., AND BUKER, S. L. Validity of the Rorschach 8-9-10 per cent. *J. consult Psychol.*, 1949, **13**, 268–271.
15. SCHOLL, R. Zur Theorie und Typologie der teilinhaltliche Beachtung von Form und Farbe. *Z. Psychol.*, 1927, **101**, 225–320.
16. SHNEIDMAN, E. Schizophrenia and the MAPS Test. *Genet. Psychol. Monogr.*, 1948, **38**, 145–223.
17. SHORR, J. A proposed system for scoring the TAT. *J. clin. psychol.*, 1948, **4**, 189–194.
18. WALLEN, R. The nature of color shock. *J. abnorm. soc. Psychol.*, 1948, **43**, 346–356.
19. WASHBURN, MARGARET F. A note on the affective values of colors. *Amer. J. Psychol.*, 1911, **22**, 114–115.
20. WEIDER, A., *et al. The Cornell Selectee Index: Short form to be used at induction, at reception, and during hospitalization.* New York: Cornell Univ. Medical Coll., 1944.

# THE STIMULUS VALUES OF THE SZONDI PICTURES

E. SZOLLOSI, D. E. LAMPHIEAR, AND H. L. BEST
THE GEORGE WASHINGTON UNIVERSITY

THE purpose of this study is to determine whether choice reactions to the photographs of the Szondi test form a sufficient basis for the rationale underlying "the dynamic thought process" [1, p. x] used as a basis for interpretation of Szondi test profiles. Statistical analysis of the internal consistency of the measuring technique is utilized to determine the validity with which it "reflects the personality as a functioning dynamic whole" [1, p. 7].

The eight factors of the test correspond to eight basic needs of the personality. If the need system of the subject is in a state of tension, the pictures are presumed to take on valence character, and the subject chooses pictures from categories corresponding to his own tension system [1, p. 26]. Test profiles are interpreted according to the numerical distribution of choices within the single need-categories. Interpretation is based on the combined number of choices from all six portrait sets in each category, and on the direction of the choices, i.e., whether pictures in the same category are chosen as like, as dislike, as like and dislike, or not chosen at all. The choice reaction of the subject to pictures of a particular type depends on the subject's unconscious attitude toward that particular need, and the number of choices indicates the strength of the need.

With such an interpretation, the Szondi test falls in the realm of projective techniques. A relatively unstructured situation is presented to the subject, who structures that situation in varying degrees in terms of his inner dynamics. Projection in the Szondi test occurs when the subject structures the stimulus material in terms of like and dislike values.

Let any quality which leads the subject to like or dislike a picture be called a stimulus value. The problem can be stated thus: Are the stimulus values among the six pictures of each disease category similar enough to justify the reasoning involved in the interpretation of test profiles? The hypothesis presented here is that the six pictures in each of the eight disease categories are unequal not only in stimulus strength but also in the quality of the stimuli and therefore evoke different choice reactions. This difference in stimulus value within categories may be of such magnitude as not to permit valid interpretation of test profiles in terms of factorial constellations (categories).

## MATERIAL

For our investigation the standard Szondi test was used [3]. It was administered in the routine way [1, p. 8]. Choices were recorded by set and category so that the pictures chosen

TABLE 1

FREQUENCIES OF CHOICE REACTIONS TO THE SZONDI PICTURES IN OUR SAMPLE OF 281 SUBJECTS

| *Set* | *h* | *s* | *e* | *hy* | *k* | *p* | *d* | *m* |
|---|---|---|---|---|---|---|---|---|
| | | | | *Like Choices* | | | | |
| I | 54 | 34 | 147 | 94 | 9 | 55 | 30 | 143 |
| II | 92 | 92 | 24 | 10 | 48 | 78 | 81 | 141 |
| III | 71 | 81 | 48 | 14 | 21 | 125 | 73 | 133 |
| IV | 74 | 164 | 40 | 41 | 80 | 117 | 30 | 20 |
| V | 206 | 43 | 43 | 43 | 56 | 70 | 23 | 82 |
| VI | 141 | 48 | 67 | 91 | 75 | 24 | 36 | 84 |
| Total | 638 | 462 | 369 | 293 | 289 | 469 | 273 | 603 |
| Mean | 106 | 77 | 62 | 49 | 48 | 78 | 46 | 101 |
| $\chi^2$ | 154 | 151 | 158 | 136 | 83 | 93 | 67 | 115 |
| | | | | *Dislike Choices* | | | | |
| I | 83 | 95 | 35 | 26 | 191 | 46 | 82 | 8 |
| II | 46 | 72 | 110 | 191 | 30 | 53 | 34 | 30 |
| III | 64 | 54 | 24 | 221 | 118 | 32 | 12 | 41 |
| IV | 57 | 30 | 69 | 54 | 64 | 27 | 151 | 114 |
| V | 9 | 68 | 109 | 61 | 70 | 72 | 92 | 85 |
| VI | 20 | 111 | 43 | 64 | 84 | 118 | 83 | 43 |
| Total | 279 | 430 | 390 | 617 | 557 | 348 | 454 | 321 |
| Mean | 47 | 72 | 65 | 103 | 93 | 58 | 76 | 54 |
| $\chi^2$ | 83 | 58 | 108 | 324 | 169 | 97 | 152 | 141 |

TABLE 2

FREQUENCIES OF CHOICE REACTIONS TO THE SZONDI PICTURES IN HARROWER'S DATA FOR 315 SUBJECTS*

| *Set* | *h* | *s* | *e* | *hy* | *k* | *p* | *d* | *m* |
|---|---|---|---|---|---|---|---|---|
| | | | | *Like Choices* | | | | |
| I | 44 | 31 | 171 | 118 | 2 | 62 | 16 | 172 |
| II | 72 | 73 | 15 | 1 | 55 | 95 | 91 | 217 |
| III | 71 | 54 | 71 | 4 | 14 | 198 | 86 | 120 |
| IV | 45 | 202 | 65 | 29 | 77 | 155 | 18 | 20 |
| V | 200 | 52 | 22 | 53 | 54 | 120 | 30 | 190 |
| VI | 142 | 29 | 80 | 64 | 144 | 31 | 26 | 90 |
| Total | 574 | 441 | 424 | 269 | 346 | 661 | 267 | 728 |
| Mean | 96 | 74 | 71 | 41 | 58 | 110 | 45 | 121 |
| | | | | *Dislike Choices* | | | | |
| I | 109 | 90 | 23 | 7 | 208 | 30 | 138 | 6 |
| II | 61 | 104 | 135 | 216 | 24 | 35 | 14 | 15 |
| III | 64 | 63 | 25 | 226 | 120 | 13 | 51 | 73 |
| IV | 106 | 22 | 42 | 83 | 41 | 14 | 162 | 117 |
| V | 32 | 51 | 135 | 73 | 82 | 68 | 93 | 56 |
| VI | 47 | 135 | 57 | 60 | 53 | 112 | 131 | 21 |
| Total | 419 | 465 | 417 | 665 | 528 | 272 | 589 | 288 |
| Mean | 69.33 | 77.5 | 69.5 | 110.82 | 88 | 45.33 | 98.16 | 48 |

*The frequencies in this table were estimated from graphs in which the data were originally presented to illustrate lectures. In reconverting the graphs, certain errors occurred of which Harrower is well aware. Statistical analysis is not justified, but the author concurs in the use of the data for comparison. Note that the frequencies of like and of dislike choices do not add up to 630 in every set.

by each subject were known. The first profile was used when the subject had been given the test more than once. The sample consisted of 283 adults between the ages of 21 and 58 and of varying nationalities and professions. All were well-functioning individuals judging by their Szondi profiles and by their overt behavior, which was determined by observation. Table 1 shows the frequency with which our sample of 283 subjects chose each picture as like and as dislike.

Toward the completion of this study, new material by Dr. Molly Harrower appeared in the Szondi Newsletter [2]. Harrower in this paper presents the choice reactions of 100 normal males, 115 females with psychological difficulties, and 100 males with psychological difficulties. Choice reactions of the 315 subjects are given by set and category in Table 2. These data will be used to support our conclusions.

## DISCUSSION

*Equality of Stimulus Strength.* According to Szondi, the 48 pictures, from a statistical point of view, represent samples from eight different, but homogeneous, universes, each sample consisting of six photographic portraits. The six pictures in any one category are supposed to measure the same thing. Therefore, they must be assumed to be equi-ambiguous as to stimulus strength in order to permit interpretation based on enumeration of choice reactions. Otherwise, interpretation would depend on pictures chosen and not only on the numerical score.

Even taking into account the highly dynamic nature of human personality, we can discount the possibility of important change in the subject's need system during the administration of the six sets. Szondi [3, p. 29] and Deri [1, p. 17] both recommend a minimum of one day between tests to detect personality changes. Furthermore, the score is evidently independent of the sets from which the pictures are chosen. Therefore, in giving the test to a large number of people sufficiently varied as to need, one would expect from the assumptions of independence and equi-ambiguity that the six pictures in a particular category would be chosen with nearly the same frequency. A comparison of frequencies from row to row in Tables 1 and 2 shows that a consistent distribution of choices from set to set is not realized in either our or Harrower's sample. The tendency toward equality cannot be expected for pictures within the same set since no assumption has been made that needs are of equal strength throughout a population or that the pictures reflect each need equally well.

To estimate the significance of the observed deviations from the column means, we apply the chi square test. Like choices may be treated independently of dislike choices since, in practice, choice reactions are well defined. Choice reactions within the same set are dependent inasmuch as once a picture is chosen, it cannot be chosen again. From set to set, choices are independent. The number of times a picture is chosen from a given category depends only on the strength of the subject's need and the number of sets

administered. The hypothesis to be tested is that sample frequencies do not vary significantly from the means of the respective categories. If this is so, the proportion of choices in any one category is the same from set to set. Then the conclusion is that choice depends on some attribute possessed in equal quantities by each picture in the category.

For our data, the $\chi^2$ values for each column are shown in Table 1. Each of these values exceeds 15.086, the maximum amount which can be attributed to chance variation for 5 degrees of freedom at the 1 per cent level of significance. In fact, the sample values are from 5 to 20 times the theoretical maximum, while an increase of only one-third improves the confidence with which we can state our results by a factor of 10. The probability of the chance occurrence of so large a value is very close to zero, but it becomes truly infinitesimal when 16 such values are found simultaneously. Therefore, we must reject the hypothesis and conclude that stimulus values are not of equal strength.

The same computation cannot be made for Harrower's data as presented in the Szondi Newsletter. However, the missing choice reactions cannot significantly alter the dispersion of choice frequencies within sets and categories, wherever they may fall. Since the dispersion of values around the column means in her data is even greater than in ours, we can expect a $\chi^2$ value of the same order of magnitude regardless of the exact distribution of the 70 missing like and dislike choices.

*Quality of Stimulus Values.* The $\chi^2$ test shows that stimulus strength varies widely among the pictures of the Szondi test. The next question that arises is whether the quality of the stimulus is the same for pictures in the same category. To find this answer, we consider a rank-order analysis.

The rank order to be considered here is based on the preference value of the picture or the algebraic sum of the plus (like) and minus (dislike) choices. From Table 3 it can be seen that the preference values range from —207 to 197 for our sample, and from —226 to 202 for Harrower's. The rank correlation between the two samples is high, with 79 per cent of the variation among choice frequencies explained by similarity of response between the two groups ($r = .89$; $r^2 = .79$). The high degree of correlation justifies the use of Harrower's data to support our findings.

TABLE 3

SZONDI PICTURES IN RANK ORDER ACCORDING TO PREFERENCE VALUE

| Picture | Our Data | | Harrower's Data | |
|---|---|---|---|---|
| | Rank Order | Pref. Value | Rank Order | Pref. Value |
| *h*V | 1 | 197 | 4 | 168 |
| *m*I | 2 | 135 | 5 | 166 |
| *s*IV | 3 | 134 | 3 | 180 |
| *h*VI | 4 | 121 | 9 | 95 |
| *e*I | 5 | 112 | 6 | 148 |
| *m*II | 6 | 111 | 1 | 202 |
| *p*III | 7 | 93 | 2 | 185 |
| *m*III | 8 | 92 | 16 | 47 |
| *p*IV | 9 | 90 | 7 | 141 |
| *hy*I | 10 | 68 | 8 | 111 |
| *d*III | 11 | 61 | 19 | 33 |
| *d*II | 12 | 47 | 11 | 77 |
| *h*II | 13 | 46 | 24 | 11 |
| *m*VI | 14 | 41 | 12 | 69 |
| *s*III | 15 | 27 | 28 | – 9 |
| *hy*VI | 16 | 27 | 26 | 4 |
| *p*II | 17 | 25 | 13 | 60 |
| *e*III | 18 | 24 | 17 | 46 |
| *e*VI | 19 | 24 | 22 | 23 |
| *s*II | 20 | 20 | 31 | – 30 |
| *k*II | 21 | 18 | 21 | 31 |
| *h*IV | 22 | 17 | 34 | – 61 |
| *k*IV | 23 | 16 | 18 | 36 |
| *p*I | 24 | 9 | 20 | 32 |
| *h*III | 25 | 7 | 25 | 7 |
| *p*V | 26 | – 2 | 15 | 53 |
| *m*V | 27 | – 3 | 14 | 53 |
| *k*VI | 28 | – 9 | 10 | 91 |
| *hy*IV | 29 | – 13 | 32 | – 54 |
| *k*V | 30 | – 14 | 30 | – 28 |
| *hy*V | 31 | – 18 | 29 | – 20 |
| *s*V | 32 | – 25 | 27 | 1 |
| *h*I | 33 | – 29 | 36 | – 65 |
| *e*IV | 34 | – 29 | 23 | 23 |
| *d*VI | 35 | – 47 | 39 | –105 |
| *d*I | 36 | – 52 | 44 | – 22 |
| *s*I | 37 | – 61 | 33 | – 59 |
| *s*VI | 38 | – 63 | 41 | –106 |
| *e*V | 39 | – 66 | 42 | –113 |
| *d*V | 40 | – 69 | 35 | – 63 |
| *e*II | 41 | – 86 | 43 | –120 |
| *p*VI | 42 | – 94 | 37 | – 81 |
| *m*IV | 43 | – 94 | 38 | – 97 |
| *k*III | 44 | – 97 | 40 | –106 |
| *d*IV | 45 | –121 | 45 | –144 |
| *hy*II | 46 | –181 | 47 | –215 |
| *k*I | 47 | –182 | 46 | –206 |
| *hy*III | 48 | –207 | 48 | –226 |

The preference values in the eight categories do not follow any pattern. For instance, in our sample, we find that *hy II* and *hy III* rank 46 and 48 out of 48 in the preference sequence, but *hy I* has rank 10, and *hy IV*, rank 16. Again, *m I* is the second preferred in our sample, while *m IV* is sixth from the last. If such variations within categories are possible, it must be that pictures were chosen not only for what they represent but for other values too. It is to be expected that both like and dislike choices will be made from the same category by any group of subjects, but there is no reason to expect certain pictures to be highly preferred while others are strongly disliked.

If *m I* was chosen 143 times as like, and *m IV*, 114 times as dislike by the same population, we have two alternatives for interpretation. According to Deri [1, p. 139], " . . . in adults minus *m* is always a sign of unhappiness," and, " . . . subjects yielding the minus *m* reactions are those who have given up hope of compensating for early frustration through new social contacts." Subjects yielding plus *m* are jolly-looking oral characters who cling to objects for the sake of oral enjoyments. According to Deri [1, p. 136], " . . . in such subjects there is a need for dependence which, if not too strong, is an asset in establishing social relations." Certainly if one of the *m*'s was liked so often for its *m*-ness, the other could not have been disliked so often for *m*-ness by the same population. At least one of the two *m* pictures was chosen for some other attribute. From this we conclude that the quality of the stimulus varies significantly within categories.

On the other hand, there can be no doubt of the correct diagnosis of illness in the photographic subjects. Most of them were taken from text books, and the rest were chosen by psychiatrists of high repute. However, anyone familiar with the limitations of diagnostic techniques in psychiatry, where the Szondi test itself may contribute to the diagnosis, realizes the difficulties in differentiating, for instance, a manic-depressive with schizoid features from a schizoid with manic-depressive features. Therefore, it should not be surprising to find stimuli in the pictures besides those connected with the principal mental disease.

To illustrate the effect of these other stimuli, let us construct a test apparatus similar in form to the Szondi test. Let there be six sets of eight kinds of objects, each well defined, from which subjects will be asked to choose in the same manner as in the Szondi test. A choice from any of the eight kinds will have a particular meaning which may lead to a Freudian interpretation.

Suppose that category 1 is defined as liquid refreshment, the choice of which measures "personality characteristics which can be derived from the early oral phase of development." Within the six sets, these refreshments vary from a glass of chlorinated water to a double shot of Haig & Haig. Category 2 contains portraits of movie actors which measure the sexual need or "need for tender love," and they vary from the portrait of Bela Lugosi to that of Rudolph Valentino. Category 3 contains books on psychology and measures the need for ego dilatation. They vary in subject matter from a study on the eating habits of the cow to Freud's *Collected Works*. In category 4, currencies measure the "anal aspect of object relationship." They vary from a one zlotyi bill to a ten thousand dollar bill.

Let Set I contain the portrait of Bela Lugosi, the ten thousand dollar bill, Freud's *Collected Works*, the glass of water, and four objects to which the subject is indifferent. Let the subject be an oral character with high need for tenderness and no overt anal needs whatever, and otherwise perfectly normal. Unless he is dying of thirst, he will select as his first like choice the ten thousand dollar bill, while in spite of, or because of, high need in the sexual vector, the Bela Lugosi portrait will be taken as the most disliked. However, this need may find a satisfactory outlet in the reading of Freud's *Collected Works*, taken as second like. He may or may not reject the water according to his reactions to the remaining objects.

For the second subject, take a Scottish farmer whose anal character needs no further description. Face him with a set containing the Haig & Haig, the study on the eating habits of the cow, the self-signed portrait of Rudolph Valentino, and a one zlotyi bill. Needless to say, the double shot of Haig &

Haig will be gulped down before the other three choices are even considered. For the same anal need, Rudolph Valentino's portrait may be chosen as second like. On the other hand, having no use for experimental and comparative psychology, he rejects the study on the eating habits of the cow. As to the one zlotyi bill, which supposedly measures his anal needs, but which is not negotiable, its chances of being selected as dislike will depend on the material values of the remaining four objects.

From these two illustrations it can be seen how choices may depend on the subject's needs, and yet these needs may not be reflected by the nominal classification of the objects chosen.

*The Forced-Choice Situation.* Another effect arises from the forced-choice situation. The subject can choose pictures from a category only once in each set regardless of the strength of the need; yet, if he is indifferent to five or six of the pictures, he must still make choices from them which are given the same weight in the interpretation of the score.

The effect of the forced-choice situation in a test where choice reactions are responses to stimuli unequal in strength and undefined in quality was not considered by Szondi. Szondi [3] believes that profiles which depart from the theoretical average are the result of reponses to stimulus values. It is not hard to see how he reached this conclusion since the pictures were carefully chosen and their classification was verified by association tests.

From the six sets, 12 pictures are chosen as like, or an average of 1.5 for each of the eight categories. Two like or dislike choices in one category are interpreted as a reflection of dynamic needs in the subject's personality, yet it is impossible to make 12 like choices among eight categories and not take two from at least one of them. Thus, we have an interpretation even though the deviation on which it is based may not be statistically significant, which shows the extreme sensitivity of the test.

The presence of a picture with high stimulus value affects choice reactions in the set regardless of the subject's needs. Before the first selection, the a priori probability that a picture will be chosen as like (or as dislike) is 1/4. If it is known that among the eight pictures, one will be chosen as like 100 per cent of the time, for example, then the a priori probabilities for selection as like become 1/7 for seven pictures, and 1 (certainty) for the picture with high stimulus value. On the other hand, the a priori probabilities for choice as dislike become 2/7 for seven pictures and zero (impossibility) for the picture which is liked 100 per cent of the time.

In our sample, it can be seen that with the exception of set V, three or four pictures are outstanding. For instance, in set I, *e* is chosen 147 times and *m* 143 times as like, while *k* is chosen 191 times as dislike. If we assume that we know a priori the frequencies of these three choice reactions in a sample of 283 subjects, then we can compute the expected frequencies of the remaining 13 choice reactions, assuming that among the remaining choices, the occurrence of each possible reaction is equally likely.

Of the 566 like choices, 147 and 143 were used up for *e* and *m*, leaving 276 choices to be distributed among the remaining six categories. Since we assume that *k* is disliked by 191 subjects, and since each subject can choose a picture only once, the maximum number of possible like choices for *k* is 92. In each of the five remaining categories the maximum number of possible like choices is 276. Then the total number of possible choices is

$$5 \times 276 + 92 = 1472.$$

The probability that a choice reaction will occur is the number of events favorable to the occurrence of that reaction, divided by the number of possible events, or

$$\frac{276}{1472} = .1875 \quad \text{for five pictures,}$$

$$\frac{92}{1472} = .0625 \quad \text{for } k.$$

Multiplying by 276, the remaining number of choices, gives the frequencies with which we expect the corresponding choice reactions to occur. Table 4 shows these expected frequencies of like and dislike choices for set I, and the observed frequencies in set I for our data.

The greater deviations from a flat profile in the observed frequencies can be attributed to uneven reactions to the remaining 13 choices. The sum of squares of deviations of the observed values from the set mean is 42,038. Of this, 34,326 is explained by the assumption that the three values are known.

TABLE 4
EXPECTED AND OBSERVED FREQUENCIES OF CHOICE REACTIONS IN SET I UNDER THE HYPOTHESIS

| Picture | *Like* | | *Dislike* | |
|---|---|---|---|---|
| | Expected | Observed | Expected | Observed |
| *h* | 51.8 | 54 | 62.8 | 83 |
| *s* | 51.8 | 34 | 62.8 | 95 |
| *e* | 146.9 | 147 | 31.1 | 35 |
| *hy* | 51.8 | 94 | 62.8 | 26 |
| *k* | 17.3 | 9 | 190.8 | 191 |
| *p* | 51.8 | 85 | 62.8 | 46 |
| *d* | 51.8 | 30 | 62.8 | 82 |
| *m* | 142.9 | 143 | 30.3 | 8 |

## SUMMARY

The Szondi test was administered to a sample of 283 normal adults, and the frequencies of choice reactions to each picture were noted.

The $\chi^2$-test applied to choice frequencies shows such wide dispersions in each category that it is virtually impossible that the frequencies are the result of reactions to pictures of equal stimulus strength.

Rank-order analysis based on the preference values of the pictures shows that choice reactions within each category span almost the entire range of variation. This combined with the contradictory interpretations attached to choice reactions within each category indicates that pictures of a given classification contain more than one quality of stimulus.

An illustration showed that pictures chosen may depend on the need system of the individual, and yet these needs may not be reflected in the classifications of the pictures selected.

It was shown that with forced choices, stimulus values as high as those contained in the Szondi pictures produce large deviations from a flat profile, regardless of the quality of the stimulus.

## CONCLUSIONS

The pictures of the Szondi test are not equal in stimulus strength, nor is the quality of stimulus the same for pictures in the same category. Therefore, interpretation based on "factorial constellations" does not reveal all the information contained in the choice reactions.

It does not follow that we should reject Szondi's ingenious idea of using pictures to reveal personality traits, but only the manner in which choice reactions are interpreted. There are two alternatives. Interpretation may be based on factorial constellations, and, in this case, pictures must be changed and rearranged until consistent reactions from set to set are obtained for any group. Or the present interpretation of the test can be dropped and reactions to individual pictures studied. Even using interpretation based on enumeration of choice reactions, the examiner often obtains more information from the spontaneous remarks of the subject than from his profile.

*Received February 14, 1951.*

## REFERENCES

1. DERI, SUSAN. *Introduction to the Szondi Test.* New York: Grune & Stratton, 1949.
2. HARROWER, MOLLY. *Szondi Newsletter.* Psychological Service Center, Univ. Oklahoma, 1950, March, 1–23.
3. SZONDI, LIPOT. *Experimentelle Triebdiagnostik.* Bern: Hans Huber, 1947.

# PSYCHOLOGICAL CASE REPORTING AT PSYCHIATRIC STAFF CONFERENCES

JULES D. HOLZBERG, SALVATORE L. ALESSI, AND MURRAY WEXLER

CONNECTICUT STATE HOSPITAL, MIDDLETOWN, CONNECTICUT

WHATEVER the nature of the psychiatric setting in which the clinical psychologist functions, he will be expected to meet with his professional colleagues in psychiatry and social work in order to discuss cases with which all or some of these workers have had contact. It is this meeting, the clinical staff conference, that has been for so long one of the basic vehicles making for the "team" approach in psychiatric practice. If this "team" approach to an individual's problems has any intrinsic worth, then the importance of this conference cannot be overstressed. While the ostensible purpose of the meeting is to permit the various professional disciplines to contribute to the understanding of the case at hand, an equally important, but not often emphasized, purpose is to provide a medium for the mutual education of professional workers. The latter purpose is as important for the psychologist as it is for his professional colleagues.

While it is obvious that participation in the clinical staff conference is an important function for the psychologist, it is of interest that little attention has been given either in universities or in clinical settings to the process of psychological case reporting at staff conferences. A cursory review of the literature revealed little if any attempt to discuss this aspect of the psychologist's function in terms of techniques of reporting and in terms of problems involved in the psychologist's participation in the staff conference.

Two techniques of reporting are prevalent today in clinical practice: (1) the psychologist reads the psychological report verbatim as it was written for the psychiatric record, or (2) the psychologist gives a generalized summary consisting entirely of interpretations. It is the feeling of the authors that neither of these two techniques satisfies the two purposes of the staff conference, i.e., clarification of the patient's status and mutual education. The first technique ignores the fact that the psychological report written for the psychiatric record may be oriented toward a goal that is distinctly different from that of a report presented at a case conference. Thus, a written psychological report may have been prepared to assist in differential diagnosis, whereas the case conference may be geared toward evaluating the most desirable therapy for the patient. In addition, the organization and style of a written psychological report does not lend itself readily to verbatim reading because of its formal character.

The second technique of reporting described above is one in which the psychologist presents abstractions regarding the patient's behavior. The abstractions may lie considerably removed from the concrete, and therefore practical, aspects of behavior in which the psychologist's professional colleagues are likely to be interested. This type of report frequently loses the individuality of the patient by dealing with broad and vague generalizations rather than with concrete data about the patient.

All psychological reporting at case conferences may not fall into these two discrete categories, and frequently a presentation may consist of both of these approaches; but the net result has often been that the colleagues of the psychologist begin to look upon his participation in the staff conference as a necessary evil, i.e., he should be there in order to give public demonstration of the "team" approach, but his reports fail to help in understanding the patient and in fostering mutual education.

The purpose of the present paper is to present a guide to more adequate psychological reporting at clinical case conferences which takes into consideration the limitations of the practices described above.

## GUIDES FOR PSYCHOLOGICAL REPORTING

*Preparation.* As indicated above, the psychological report written for the psychiatric record is frequently not in a desirable form for presentation at case conferences. Consequently, the psychologist must prepare specifically for the case conference. When the patient has been seen by the psychologist weeks or possibly months before the conference, it is especially important that the psychologist review the case and prepare an outline of the material that he will present. The frequent practice of reaching for the psychological documents a few minutes before the conference and attempting to ad-lib a report can only result in a very poor presentation.

Preparation is vital to the psychologist's participation at the staff conference. This preparation involves the development of an outline of the relevant material to be discussed. However, such an outline, once prepared, should not be used rigidly. It may be necessary to reduce the length of presentation because of limitations of time, or it may be important to focus more intensively on fewer factors. Here, there can be no substitute for the psychologist's utilizing good judgment in determining the use of any outline that he prepares.

Frequently, the psychologist looks upon the case conference as a burden and would prefer to devote his time to other psychological activities. This attitude is unfortunate in that the psychologist fails to recognize the importance of the case conference from the point of view of the patient, the psychologist, and psychology as a profession. The case conference should be accepted as an integral part of the total responsibility of the psychologist.

*Pertinence.* The psychological report prepared for the psychiatric record should contain as much detail as possible since what appears as a minor factor at the moment may take on greater significance at some future time. Consequently, it is not necessary to present at the case conference everything that appears in the formal psychological report.

A case conference usually has a specific goal in mind. A conference may be called in order to establish a nosological or etiological diagnosis, and here the psychologist should be concerned with contributing maximally to this goal. Another conference may be called to discuss the therapeutic possibilities for a given patient, and here the emphasis of the psychologist should be geared toward meeting this objective. Another type of conference may be concerned with the desirability of discharge of a patient from a mental hospital. This may require a different emphasis in the psychologist's presentation from either a diagnostic or therapeutic conference. A conference may be called with reference to the disposition of a court case. Here the psychologist should contribute material that has pertinence for the purpose of the conference. Thus, with the case conference concerned with specific problems, the psychologist is responsible for preparing his material so as to contribute maximally to the goal of the conference.

Here an objection may be raised by the psychologist who feels that what he may omit from his presentation may nevertheless be very vital for understanding the patient. This necessitates consideration of another aspect of the psychologist's participation at the staff conference. The role of the psychologist at the conference may be considered in two phases: (1) the formal report based on an outline prepared in advance of the conference which draws together data pertinent to the patient's particular problem being discussed; (2) the psychologist's participation in the discussion of the case subsequent to the presentation of the formal report. The second aspect of the psychologist's role in the case conference is often underemphasized in clinical practice. Frequently, the psychologist will look upon his role as merely that of presenting a formal report and then withdrawing from discussion in subsequent stages of the conference. At these latter stages the psychologist can introduce other material about the patient which at that particular moment becomes pertinent to the staff discussion. Parenthetically, it should be noted that at this stage of the discussion the psychologist should not feel himself confined to his test data. He should feel free to make contributions concerning the problem at hand even though his ideas are not derived from his tests.

In some settings, the other participants in the staff conference share responsibility for the psychologist's failure to accept the second phase of his role in the case conference. Thus, it has been observed that the psychiatrist frequently accepts the psychologist's role as simply that of preparing a formal report without encouraging his participation in the subsequent discussion.

The tendency of the psychologist to feel that his contribution is finished once he has delivered his formal report may have serious consequences for the patient. The psychiatrist, in attempting to integrate the psychologist's thinking into the total psychiatric picture, may misunderstand the psychologist's reasoning, resulting in distorted conclusions about the patient. For example, a psychologist reported at a staff conference that a patient was functioning in the defective range of intelligence. He reasoned that there was evidence to indicate that certain morbid processes were seriously disturbing intellectual activity and that the intellectual potential was in all probability considerably higher. The psychiatrist, in summarizing the case, overlooked the psychologist's evaluation. The former stated that because of the defective intelligence, there was little likelihood that the patient would benefit from psychotherapy. If the psychologist in this case had not further elaborated upon his interpretation of the patient's performance, a disservice might have resulted.

*Organization.* The psychologist's report at the staff conference should be well organized and integrated. While this may appear to be an obvious principle requiring no special emphasis, it is a principle honored more often in the breach than in the observance. If the purpose of the conference is to answer a specific problem, the psychologist should arrive at his solution through a process of logical reasoning. In order to achieve this, the psychologist should not report on a test-by-test analysis, but should make a cross-test analysis of the pertinent psychological factors.

*Evidence.* The inclusion of verbatim test material in the presentation of the psychological report is vital if the dual purpose of clarifying the patient's problems, and mutual education is to be served. If one of the purposes of the staff conference is to provide better understanding of a patient, then the goal is to present a rich and vibrant picture of the personality. If the psychologist compounds his conference reports with a series of psychological clichés and poorly deduced generalizations, then his auditors at the staff conference will soon begin to feel that tests are unable to discriminate between different personalities. As one moves from the test data and its specific interpretations to broad dynamic formulations, the patient, as a specific personality, begins to lose his identity. These remarks are not to be construed as opposed to the presentation of generalizations that are supported by specific and concrete data.

An additional reason for the inclusion of test data is to fulfill the second purpose of the staff conference, namely, mutual education. Perhaps nothing is more important in the process of mutual education than the sharing with other professional personnel of the thinking that went into the analysis of a given case, beginning with the original test data.

For example, the psychologist may report, without any preliminary attempt to build up or present the evidence, that the patient is "an orally sadistic individual." In capsule form, the psychologist is attempting to convey a motivational formulation, but it loses its richness because the specific ways in which the patient thinks and feels about this problem are absent.

For presentation at the staff conference, such a formulation will become meaningful and instructive to the other participants if the psychologist presents his data, and his ways of reasoning from his information. Instead of the generalization, "an orally sadistic individual," the following method of presentation is suggested. "The Rorschach data reveals several percepts which suggest that sadism is related to the oral zone. 'A crocodile snapping his jaws shut,' 'a man and woman arguing,' 'a snake poised to sink his fangs into his victim.' Other tests such as word association test and the Thematic Apperception Test offers evidence which further demonstrates the oral sadistic trends. To the word 'kiss,' the response is 'bite'; to the word 'talk,' the response is 'argue.' On one card of the Thematic Apperception Test (7 BM), the patient says, 'These are two politicians. The young one has just debated with his opponents. His words were so brilliant and *stinging* that he literally *destroyed* all opposition.'" The evidence that went into the conclusion "oral sadism" is presented. It gives the listener an opportunity to witness the patient's

thinking. The interpretive job is by no means complete with the observation that the patient is orally sadistic. "Oral sadism" telescopes a great many ideas, and simply to attribute such traits to a patient does not convey much that is meaningful about him. Thus, the following is a continuation of the psychologist's description of the patient. "The data suggest (kiss—bite) that the erotic relationships seem to be colored by sadistic traits. Further, argumentation may be a vehicle through which these traits may be expressed. Words seem to have a magical significance and are used in an aggressive and destructive manner. The Thematic Apperception Test data concerning oral sadism and the psychological importance of language to the patient can be further illustrated by responses on the Wechsler-Bellevue. For example, to the word 'hara-kiri,' the patient says, 'This is an ancient and honored custom among the Japanese. It derives from an innate sense of honor to the Motherland and from a philosophy of Stoicism which only the oriental people have. It consists of driving a blade into the stomach and ripping out the viscera. I find a certain fascination in contemplating the scene.'" Such examples clarify and elaborate the psychologist's reasoning and make for a richness of reporting which will both hold the listener's interest and instruct at the same time. The Chinese proverb that one picture is worth a multitude of words is as true in psychology and psychiatry as it is in other areas of instruction.

Occasionally, the psychologist will offer objection to the idea that he should present evidence in substantiation of his interpretations and generalizations. He may consider this a reflection on his professional status and will correctly raise the question, "Do the other professional workers offer evidence when they are presenting cases?" The answer to this objection is twofold. (1) There is no inconsistency between presenting test data as evidence for one's thinking, and maintaining one's professional status. In reality, it is an earmark of a scientist carefully at work, who is determined to make no interpretations and generalizations without adequate evidence. (2) The failure of other professional groups such as psychiatrists and social workers to offer evidence to support their thinking should be as readily condemned as the psychologist's failure to do so. The psychiatrist who offers a diagnosis or therapeutic plan without any attempt to build up the evidence for his thinking carefully, and the social worker who gives a judgment regarding a family situation without presenting the underlying evidence, are engaging in unscientific practices that can only lead to failure to attain the purposes of the staff conference.

*Content.* There is certain minimal information that should be incorporated into the psychologist's presentation of a case at staff conference. The following sequence is not meant to suggest the order of presentation at the staff conference, although for most cases this order should prove satisfactory.

1. Date of examination. This is important because a patient may have been last seen several weeks and possibly months before. Significant psychological changes may have occurred since the patient was seen. Knowledge on the part of the other staff members of when the psychologist saw the patient places the report in proper perspective.

2. Reason for referral. This defines the responsibility of the psychologist in terms of the type of psychological study he has performed. It may be stated with reasonable assurance that most clinical settings cannot afford to engage in extensive testing of patients but must confine testing to the exploration of specific problems that have been raised. Consequently, a given patient may not have been studied for a problem which suddenly emerges at the staff conference. A patient may have been seen by the psychologist in order to explore the personality of the patient to aid in defining therapeutic objectives, and the entire battery of psychological tests may have been directed toward this end. At the staff conference, however, a problem may arise concerning vocational rehabilitation which the psychologist had not explored with the patient. Thus, it is important at the outset of the report to structure carefully the reasons why the patient was referred for psychological study.

3. Behavioral observations. This aspect of the psychologist's report to the staff conference is frequently underemphasized for two reasons: (*a*) behavioral observations by the psychologist approach in content the mental status examination performed by the psychiatrist. The psychiatrist may therefore look upon the reporting of behavioral observations by the psychologist as an intrusion into his domain of study; (*b*) the psychologist may fail to recognize that the study of the patient's behavior while taking tests may be as important as the data procured from the tests themselves. It is the authors' contention that these two reasons for the psychologist's neglecting to report behavioral observations are totally invalid. If one views the administration of the psychological examination as an opportunity to observe the patient under standardized interview conditions, such observations should be an important supplement to the psychiatrist's observations derived from relatively unstandardized interviews. From the psychologist's point of view, the behavior

of the patient defines the frame of reference within which the interpretation of test data can be made. If the psychologist is to be "patient-oriented" rather than "test-oriented," the reporting of behavioral observations is a vital part of his contribution at the staff conference. The patient's attitude toward the examiner, toward the pictures on the wall, or toward the challenge of being tested may be just as significant as his formal responses to a specific test.

4. Intelligence. Where an intelligence test has been given, a statement of the patient's present functioning level and his optimal capacity should be reported even if no further report on the intelligence test results will be made. This type of information, just as behavioral observations, is vital for structuring the frame of reference within which to understand the patient's performance on other tests.

5. The main report. The body of the report is concerned with the concrete evidence, the interpretations, and the generalizations. When previous psychological studies have been performed, these should be incorporated for comparison purposes.

6. Conclusions. The psychologist should complete his report, rendering his opinion regarding the problem for which the patient is being presented. It should be an integration of all that the psychologist has previously reported. However, the summary should not merely be a repetition. The psychologists's conclusions should be qualified in terms of the degree of certainty with which he offers his opinion. His conclusions may be tentative hypotheses, and then should be stated as such. On the other hand, the conclusions may be definitive and supported by voluminous evidence, in which case the psychologist should voice the certainty of his opinion. It must be emphasized that a conclusion regarding human behavior at our present state of knowledge permits varying degrees of assurance. It is no more a reflection on the psychologist to qualify his opinions than it is on his colleagues in other professions to qualify theirs.

## MANNER

While it is not the authors' intention to urge a Dale Carnegie approach to psychological reporting at the staff conference, it must, however, be stressed that unless the psychologist is enthusiastic and interested in what he is presenting, there is not likely to be much enthusiasm or interest on the part of those that are listening. It is amazing to discover the number of psychologists reaching advanced stages of training in clinical psychology who are unable to meet with colleagues and to exchange thinking without experiencing feelings of severe threat and debilitating anxiety. The net result of these feelings is to prevent intelligent and effective participation in the case conference.

The use of technical psychological language does not make others accept the psychologist as a scientifically trained person. It should be stressed again that his status as a scientific worker will depend upon his ability to present evidence logically and convincingly in support of his interpretations and generalizations. The injudicious and unexplained use of such terms as "extratensive," "flexor," "deterioration ratio," "succorance need," etc., while ringing with scientific erudition, will only result in isolating the psychologist from his professional colleagues.

This paper has stressed that a concrete approach to the staff conference is most desirable. Concreteness can be emphasized most strongly when the psychologist comes to the staff conference prepared to present the actual test materials with which the patient has worked. Thus, presenting the Rorschach card on which the patient has given an unusual response, the TAT card which has produced a significant story, the human-form drawing which takes on special significance in the psychologist's analysis of test data, or the Wechsler-Bellevue subtest that delineates the special problem area in the study of the patient—these and any other concrete presentations of test material should make for a richness in psychological reporting that will not be equalled by the mere presentation of interpretations and generalizations.

## CONCLUSION AND SUMMARY

This paper has attempted to focus the need for clinical psychologists to receive training in the proper techniques of reporting at psychiatric staff conferences. Principles have been enumerated which are considered appropriate for the presentation of psychological reports. While it is not the authors' intention to submit these principles as the last word on the problem, it is hoped that they may stimulate thinking about the psychologist's role and function at the staff conference, based on a recognition of its importance in his daily professional life.

*Received January 5, 1951.*

# THE EFFECT OF BRIEF STIMULUS ELECTROCONVULSIVE THERAPY UPON PSYCHOLOGICAL TEST PERFORMANCES[1]

ISIDOR W. SCHERER[2]

VETERANS ADMINISTRATION HOSPITAL, NORTHAMPTON, MASSACHUSETTS

THIS study examines the effects of electroconvulsive shock using the Liberson brief stimulus therapy (BST) technique, [3] which employs a square wave of very low pulse duration. The effects were measured by a battery of psychological tests administered to mental patients, mainly in a group situation.

A majority of workers in the field of clinical psychology are aware of the need for the objective study and evaluation of the numerous neuropsychiatric treatment procedures applied today. They realize that unless valid comparative studies are made between the various types of treatment, such as individual and group psychotherapy, activity therapy, the various shock therapies—insulin coma, insulin sub-coma, metrazol, electric shock, and lobotomy—it will become increasingly difficult to select with any degree of assurance a particular type of treatment for the various mental disease syndromes. Of further concern is the belief that some treatment techniques produce harmful organic results in areas of intellectual function which may be obscured by the apparent pyschosocial gains claimed and may thus, in a sense, nullify the partial benefits of the treatment. The present report is concerned with an attempt to investigate possible intellectual and emotional damage resulting from one of the brief stimulus electroshock therapy procedures.

[1]One of four experiments submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy at the University of Nebraska. Abstract of this paper was presented to Psi Chi at the Amer. Psychol. Assn. meeting, Boston, September, 1948.

[2]Chief Clinical Psychologist, Veterans Administration, Northampton, Mass. Published with the permission of the Chief Medical Director, Department of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed or conclusions drawn by the author.

[3]The Liberson Brief Stimulus Electroconvulsive Therapy Apparatus, Type 735, manufactured by Offner Electronics, Inc., Chicago, was used in this experiment.

## GENERAL PROCEDURE

*Choice of Tests.* Numerous writers, including Brody [2], Levy *et al.* [6], Bennett [1], Lowenbach and Stainbrock [10], Alpers and Hughes (reviewed by Lewis [7]), and Ziskind *et al.* [11] claim that electric shock treatment causes intellectual impairment and brain damage. Liberson [7, 8, 9], on the other hand, contends that the main advantage of brief stimulus electroconvulsive therapy is the avoidance of memory disorders and other intellectual impairment. In the present experiment we have applied certain memory tests to check these divergent claims. Tests included were: immediate and delayed recall, distortion per cent, and memory efficiency from the Babcock Memory Paragraph; Benton Visual Memory; Hunt-Minnesota Hard Pairs (auditory); and per cent recall on the Word Association Test. If it were possible to capture the subjective, but significant, estimates of memory ability, our evaluations might be more satisfactory. At the present time, however, we are unable to do this, and, therefore, we must rely upon tests which may not be related to some of the memory disabilities which are reported by the physicians and by the patient's relatives and friends.

In addition to tests of memory function, abstraction or concept formation tests were administered in order to detect possible brain damage resulting from the treatment. Tests

in this category included the Shipley–Hartford Retreat Scale, Goldstein's Object Sorting Test, the Similarities subtest from the Wechsler-Bellevue Scale and Capp's Categorization Test. A precedent for the use of such tests is the point of view discussed by Goldstein and Scheerer [4, p. 9], who state, "Organic pathology in patients with brain disease disintegrated human behavior in such a way that the capacity for abstract behavior is impaired to a greater or lesser degree in the patient." The findings of Greenblatt *et al.* [5, p. 390] are especially interesting. On the basis of a comparative study between the value of electroencephalography and abstractive tests for the diagnosis of brain damage they concluded that, "Abstraction tests were more reliable in the diagnoses of brain damage in this heterogeneous sample of cases than were the electroencephalograms."

Motor tests including Minnesota Rate of Manipulation and the Digit Symbol subtest of the Wechsler-Bellevue were administered for the following reasons:

1. In the beginning of the experiment the electrodes were placed in the vertex region, very close to the giant Betz cell area. It was hoped that by using tests of gross and fine motor coordination, it would be possible to discover any motor reactivity changes which might result from direct stimulation of this motor area.
2. To obtain a measure of psychomotor speed and learning capacity for simple tasks.
3. To explore the possibilities in using the motor function modality for measures of cooperation. The writer feels that such tests offer simple, reliable estimates of at least superficial cooperative levels. It is becoming increasingly apparent that the patient's cooperation plays an important role in test achievement as well as in his improvement.

The Rorschach Multiple Choice Test was administered in order to observe any changes which occurred in personality structure following BST. It was thought that it might be possible to find significant response choices which would be reliable prognostic indicators.

The Vocabulary test was used mainly to obtain a memory efficiency index, and also to obtain a crude measure of intelligence and a qualitative pattern analysis.

The Digit Span test was used as a measure of attention. The reverse Digit Span score was used as another measure of organic brain damage, since there are numerous references concerning this relationship in the clinical literature.

*Subjects.* The experimental group (E) consisted of 41 mental patients[4] who were given a complete course of electroshock treatments. These ranged from 10 to 31 treatments, with a mode of 20. BST was terminated when the medical staff felt improvement was fixed or that the patient could not benefit from further treatment. These patients had an average age of 30.8 years, had attained an average of 10.97 grades in school, and had been hospitalized for an average of one year, four months.

The control group (C) was composed of 17 mental patients with an average age of 33.5 years, 10.26 grades education, and had been hospitalized for two years. The E and C groups are fairly well equated. The patients in group C had originally been chosen by the medical staff to receive shock treatment, but because of such complicating factors as physical contraindication and family refusal to permit shock treatment, they did not receive BST but were used as controls in this experiment.

*Testing Procedure.* Approximately three weeks prior to receiving BST,[5] 40 psychotic patients and 1 severe psychoneurotic were given, in a group situation,[6] the battery of

[4]At the start of the experiment the total number of patients included in the E and C groups was 165. All 165 patients were given Battery I. However, acute periods of disturbance during the required testing period, development of physical complications during the shock treatments, family interference in regard to continuation of treatment, and various factors greatly reduced the number of patients. The subjects were patients at the Veterans Administration Hospital, Northampton, Massachusetts. Of the 41 patients in the E group, 14 were diagnosed as Dementia Praecox, Paranoid; 10 as Dementia Praecox, Catatonic; 2 Dementia Praecox, Hebephrenic; 2 Dementia Praecox, Simple; 5 Dementia Praecox, Unclassified; 6 Manic Depressive, Depressed; 1 Involutional Melancholia; and 1 Psychoneurotic.

[5]In cases where vertex temporal approach was used the average peak voltage was 355, duration in seconds 1.1, and pulse duration milliseconds, .50. Bitemporal approach peak voltage was 534, duration in seconds, 1.3, and pulse duration milliseconds, .70.

[6]All tests were administered as group tests except Digit Span, Word Pairs, and Minnesota Rate of Manipulation.

TABLE 1

COMPARISON OF CHANGES IN MEMORY TEST SCORES FROM BATTERY I TO BATTERY II FOR EXPERIMENTAL AND CONTROL GROUPS

| Test | Gr. | N | Mean Score Bat. I | Mean Score Bat. II | Diff |
|---|---|---|---|---|---|
| Babcock Par. Mem. | | | | | |
| Immed. Recall | E | 40 | 9.63 | 11.38 | 1.75 |
| | C | 16 | 8.00 | 9.50 | 1.50 |
| Delayed Recall | E | 40 | 8.65 | 10.18 | 1.53 |
| | C | 15 | 6.30 | 9.00 | 2.70 |
| % Distortion, Delayed Recall | E | 39 | 19.12 | 9.25 | – 9.87 |
| | C | 14 | 33.71 | 16.64 | –17.07 |
| Diff. between Immed. and Delayed Mem. | E | 40 | – .95 | 1.20 | – .25 |
| | C | 15 | 1.73 | – .93 | .80 |
| Mem. Effic. Index | E | 37 | .86 | 1.17 | .31 |
| | C | 11 | 1.09 | 1.03 | – .06 |
| Benton Vis. Mem. | E | 40 | 4.45 | 4.50 | – .05 |
| | C | 16 | 4.56 | 4.56 | .00 |
| Hunt-Minn. | | | | | |
| Aud. Mem. Word Pairs | E | 29 | 6.91 | 8.70 | 1.79* |
| | C | 16 | 5.94 | 5.31 | – .63 |
| Word Asso. Recall % | E | 32 | 77.81 | 78.75 | .94 |
| | C | 12 | 66.33 | 63.58 | – 2.75 |

*Net change significant at 10 per cent level.

TABLE 2

COMPARISON OF THE CHANGES IN ABSTRACTION TESTS FROM BATTERY I TO BATTERY II FOR EXPERIMENTAL AND CONTROL GROUPS

| Test | Gr. | N | Mean Score Bat. I | Mean Score Bat. II | Diff |
|---|---|---|---|---|---|
| Shipley-Hartford | E | 40 | 17.20 | 18.45 | 1.25 |
| | C | 17 | 12.50 | 14.00 | 1.50 |
| Object Sorting | E | 39 | 17.79 | 19.87 | 2.08 |
| | C | 14 | 12.50 | 15.20 | 2.70 |
| Wechsler Similarities | | | | | |
| Weighted Score | E | 39 | 8.87 | 9.28 | .41 |
| | C | 17 | 7.64 | 7.76 | .12 |
| Abstraction | E | 40 | 3.10 | 4.05 | .95 |
| | C | 17 | 3.65 | 4.41 | .76 |
| Categorization | E | 39 | 21.18 | 20.43 | – .75 |
| | C | 16 | 17.13 | 18.44 | 1.31 |

TABLE 3

COMPARISON OF THE CHANGES IN RORSCHACH MULTIPLE CHOICE RESPONSES FROM BATTERY I TO BATTERY II FOR EXPERIMENTAL AND CONTROL GROUPS

| Measure | Gr. | N | Mean Score Bat. I | Mean Score Bat.II | Diff |
|---|---|---|---|---|---|
| No. Poor Responses | E | 41 | 2.12 | 4.53 | 2.41 |
| | C | 16 | 1.53 | 3.72 | 2.19 |
| Movement, Crude Measure | E | 41 | .68 | .85 | .17 |
| | C | 16 | .75 | .69 | – .06 |
| Poor Form, Crude Measure | E | 40 | .70 | .35 | – .35† |
| | C | 16 | .06 | .75 | .69 |
| Rejections | E | 41 | 1.14 | 1.68 | .54 |
| | C | 16 | 1.81 | 1.50 | – .31 |
| Color, Crude Measure | E | 41 | 1.24 | 1.07 | – .17* |
| | C | 16 | .87 | 1.50 | .63 |

†Net change significant at 5 per cent level.
*Net change significant at 10 per cent level.

TABLE 4

COMPARISON OF THE CHANGES IN MOTOR, ATTENTION, AND VOCABULARY TESTS FROM BATTERY I TO BATTERY II FOR EXPERIMENTAL AND CONTROL GROUPS

| Test | Gr. | N | Mean Score Bat. I | Mean Score Bat. II | Diff |
|---|---|---|---|---|---|
| Motor | | | | | |
| Rate of Manipulation Test | E | 38 | 82.34 | 80.42 | –1.92 |
| | C | 15 | 84.93 | 80.40 | –4.53 |
| Wechsler Digit Symbol | E | 41 | 8.68 | 8.07 | – .61 |
| | C | 16 | 8.40 | 7.67 | – .73 |
| Attention | | | | | |
| Wechsler Digit Span | | | | | |
| Total | E | 41 | 8.10 | 8.10 | .00 |
| | C | 15 | 6.87 | 7.40 | .53 |
| Forward | E | 39 | 6.59 | 6.38 | – .21 |
| | C | 15 | 6.40 | 6.67 | .27 |
| Reverse | E | 39 | 4.08 | 4.46 | .38 |
| | C | 15 | 3.33 | 3.86 | .53 |
| Vocabulary | | | | | |
| Wechsler Vocabulary | E | 41 | 9.83 | 9.78 | – .05 |
| | C | 17 | 8.65 | 8.65 | .00 |

psychological tests described above which will be referred to as the preshock battery or Battery I. The size of the group varied from 2 to 7 patients. From two to six weeks after the termination of ECT, the same group of patients was given a postshock test battery, Battery II. The tests in Battery II consisted with a few exceptions of alternate forms of Battery I.

Batteries I and II were administered to a control group of 17 psychotic patients who did *not* receive (BST) electric shock treatment. The time interval between the administration of the two batteries was approximately two months.

With the exception of several acutely disturbed patients, both the E and the C groups were housed in the reception service.

The activity and social programs were the same for both groups; however, the E group attended a group psychotherapy session at least once a week whereas the control group did not.

## RESULTS AND DISCUSSION

The results for the memory tests are summarized in Table 1, and the results for the abstraction and Rorschach tests are given in Tables 2 and 3, respectively. Table 4 presents the relevant material for the remaining tests of the battery. It may be observed that both the experimental and control groups showed evidence of improvement, due possibly to improved rapport, practice, and other factors; nevertheless, the degree of improvement from first to second testing is not greater for one group than for the other. This conclusion is based on a statistical examination of the significance of the difference between differences; the method of analysis employed is the one suggested by Edwards [3, p. 182]. Thus, the present study provides no striking or consistent evidence in support of the hypothesis that electroconvulsive therapy affects psychological performance adversely.

With respect to the memory tests, the investigator does not feel that the tests employed here provide a basis for exhaustive evaluation of the effect of the shock therapy on memory, since memory is coming more and more to connote a certain continuum of subjective meaningfulness in the individual being tested. He does feel, however, that since he has measured the effect of shock on a specific item over a specific period of time in a standardized situation, the results obtained are probably more reliable than those obtained by anecdotal methods and certain biased clinical judgments. There have appeared in the literature several reviews which claim that electroconvulsive therapy produces memory impairment and brain damage. These reviews are based upon the subjective type of analysis. One study [2] which is widely cited concluded on the basis of a subjective estimate of five patients that prolonged memory defects followed electric shock therapy. This report stated that the patients could not remember the names of people, the addresses of relatives, the contents of recipes, etc. The evidence cited does not seem to substantiate such a positive general statement, as, "Finally, it implies permanent or semipermanent damage to the brain" [2, p. 779].

Our experiment compares postshock performance with the baseline of the functional level achieved in the preshock battery. However, if some hypothetical premorbid potential above the preshock level is assumed on the basis of certain qualitative aspects of the preshock battery and case history material, then this lack of improvement as between post- and preshock results may indicate that shock has injured the patient to the extent that he is unable to achieve his premorbid intellectual potentialities, even though he can shake off the intellectually debilitating effects of the psychosis, as evidenced by positive emotional changes found on some of the tests and clinical estimates of emotional improvement. This is a general statement of the results, however, and there are particular features of the tests employed and the data they provide which require consideration and comment. Thus, the patients who underwent BST showed a definite positive change on one test item, and showed an interesting positive trend on two other items; Rorschach poor form perception was reduced ($P = .05$), suggesting increased reality contact; there was a decrease in Rorschach color-form responses ($P = .10$), suggesting a drop in regressed affectivity; and third, word pair memory ability increased ($P = .10$), suggesting increased attentive capacity. If the experimental group enjoys a general beneficial result from the shock treatment, a more reality-bound, less labile emotional level, and if the generally facilitating effect upon intellectual performance of such therapeutic progress is acknowledged, then, in view of the absence of improved memory test results on the postshock batteries, we may hypothesize the possible deleterious effects of BST on memory processes.

## SUMMARY AND CONCLUSIONS

It was the purpose of this experiment to measure the effect of brief stimulus electroconvulsive therapy on the test performance of mental patients. The brief stimulus therapy technique was chosen because its proponents

claimed a minimal amount of brain damage resulting from this procedure. A battery of psychological tests was administered before and after shock treatment to 41 patients who received electroconvulsive therapy, and the same test battery was given to a control group of 17 patients who did not receive brief stimulus therapy. The significance of the changes in the test performance of the patients was determined by a *t*-test technique. The results indicate that BST does not adversely affect performance on tests of memory, abstractive capacity, attention and concentration, motor ability, vocabulary, or the personality measures applied in this study. There is both clinical evidence and scattered test evidence that brief stimulus therapy shows some positive orientation changes. These positive effects could possibly obscure slight deteriorative physiological effects resulting from shock therapy. With this reservation, brief stimulus therapy seems to result in no discernible psychological loss measurable by the battery employed.

*Received January 5, 1951.*

## References

1. Bennett, A. E. An evaluation of the "shock" therapies. *Psychiat. Quart.*, 1945, **19**, 465–477.
2. Brody, M. B. Prolonged memory defects following electrotherapy. *J. ment. Sci.*, 1944, **90**, 777–779.
3. Edwards, A. *Statistical analysis for students in psychology.* New York: Rinehart, 1946.
4. Goldstein, K., and Scheerer, M. Abstract and concrete behavior: an experimental study with special tests. *Psychol. Monogr.*, 1941, **53**, No. 2.
5. Greenblatt, M., Levin, S., and Atwell, C. Comparative value of electroencephalogram and abstraction tests in the diagnosis of brain damage. *J. nerv. ment. Dis.*, 1945, **102**, 383–391.
6. Levy, N. A., Serota, H. M., and Grinker, R. Disturbances in brain function following convulsive shock therapy. *Arch. Neurol. Psychiat., Chicago.*, 1942, **47**, 1009–1029.
7. Lewis, N. D. Shock therapy: present status. *Bull. N.Y. Acad. Med.*, 1943, **19**, 227–244.
8. Liberson, W. T. Brief stimulus therapy: physiological and clinical observations. Paper read at the Amer. Psychiat. Ass., New York, N.Y., May, 1947.
9. Liberson, W. T. New possibilities in electric convulsive therapy "brief stimuli" technique. Preliminary Report. *Dig. Neurol. Psychiat.*, 1944, **12**, 368–369.
10. Liberson, W. T., and Wilcox, P. N. Electric convulsive therapy: comparison of "brief stimuli technique" with Friedman-Wilcox-Reiter technique. *Dig. Neurol. Psychiat.*, 1945, **13**, 292–302.
11. Lowenbach, H., and Stainbrook, E. Observations on mental patients after electroshock. *Amer. J. Psychiat.*, 1942, **98**, 828–833.
12. Ziskind, E., Somerfield-Ziskind, E., and Ziskind, L. Metrazol and electric convulsive therapy of the affective psychoses: controlled series of observations covering a period of five years. *Arch. Neurol. Psychiat., Chicago*, 1945, **53**, 212–217.

# NORMS, RELIABILITY, AND FORMS OF THE MMPI

A. R. GILLILAND AND RUSSELL COLGIN
NORTHWESTERN UNIVERSITY

THE significance of the scores on any test is dependent, among other things, upon the norms for the test. The published norms for the nine components of the MMPI were derived by giving the tests to 1,729 persons. Nearly half of this group were outpatients in the University of Minnesota Hospital. The others were distributed among high school graduates seeking admission to college, college students, patients in general wards of the university hospital, and 221 psychopathic patients. This seems to be a typical cross section of the general population except that there was somewhat more than the usual number of psychopathic patients.

Scores on the nine factors of the test were computed from the raw scores derived from the answers to the 550 items of the test. These raw scores were transmuted into standard scores with 50 as the mean and a sigma of 10. A score of 70 or above is assumed to be an indication of abnormality. The manual states, "Clearly normal persons do not often score above 70." A score of 30 or below indicates a similar deviation from the mean, but little is said about the significance of low scores.

## THE NORMS FOR THE MMPI

One purpose of the present study is to determine how typical the norms for the MMPI are, especially for groups of college students. Several studies have been reported in which the MMPI has been used with college students. Means are presented in Table 1 for various college groups at three different universities. The Minnesota groups

TABLE 1
MEAN SCORES FOR DIFFERENT COLLEGE GROUPS

| Investigator | Subjects | Scales *Hs* | *D* | *Hy* | *Pd* | *Mf* | *Pa* | *Pt* | *Sc* | *Ma* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | *University of Minnesota* | | | | | |
| Torrens [3] | 366 Females* | 46.7 | 49.9 | 53.0 | 53.8 | 53.0 | 52.2 | 48.2 | 49.2 | 54.7 |
| | 82 Males* | 43.7 | 50.0 | 53.6 | 48.3 | 55.1 | 57.3 | 46.3 | 45.8 | 49.7 |
| Brown [2] | 176 Males | 53.7 | 54.9 | 56.8 | 54.9 | 56.1 | 53.6 | 52.4 | 52.8 | 58.7 |
| | 366 Females | 46.7 | 49.9 | 53.0 | 53.5 | 53.0 | 52.2 | 48.2 | 49.2 | 54.7 |
| | 608 Males | 51.5 | 52.0 | 54.6 | 57.0 | 59.2 | 52.2 | 57.0 | 58.2 | 58.5 |
| | 324 Males | 50.7 | 52.2 | 53.7 | 58.3 | 55.1 | 53.1 | 58.3 | 57.6 | 59.8 |
| | 508 Females | 48.9 | 49.8 | 53.4 | 53.9 | 48.1 | 53.8 | 54.6 | 54.5 | 54.8 |
| | | | | | *University of Wisconsin* | | | | | |
| Blum [1] | 25 Educ. (Males) | 45.9 | 48.7 | 51.5 | 48.9 | 57.7 | 51.1 | 49.7 | 44.7 | 55.1 |
| | 25 Eng. " | 46.4 | 53.2 | 56.8 | 47.2 | 61.6 | 54.9 | 51.0 | 44.8 | 54.0 |
| | 25 Jour. " | 49.2 | 51.2 | 54.0 | 48.4 | 59.6 | 51.4 | 50.0 | 49.5 | 49.2 |
| | 25 Med. " | 50.0 | 51.8 | 54.9 | 50.3 | 58.7 | 49.9 | 51.4 | 52.8 | 49.5 |
| | 25 Law " | 49.0 | 47.8 | 55.4 | 47.3 | 55.2 | 51.2 | 49.5 | 46.8 | 52.1 |
| | | | | | *Northwestern University* | | | | | |
| Gilliland and Colgin | 160 Undergrads. | 49.4 | 51.6 | 53.6 | 54.5 | — | 52.6 | 54.6 | 56.6 | 57.5 |
| | 600 Total | 50.1 | 52.3 | 55.7 | 56.0 | 64.1 Males<br>49.9 Females | 52.5 | 60.4 | 56.7 | 57.9 |

*Not corrected for K factor.

were samples from different colleges on the campus. The Wisconsin students were 25 male students from each of five colleges, and the Northwestern group were 160 students in the beginning psychology courses, generally freshmen and sophomores from the different colleges on the campus. The total also included an additional 440 junior, senior, and graduate students, mostly majors in psychology, education, or speech.

From an examination of these data it is apparent that the mean scores in some components of the scale are consistently and significantly above 50. The component in which there was the largest and most consistent deviation above 50 was in the Interest, i. e., the Masculinity-Feminity score for men. In all nine groups where men and women were segregated, the mean for men varied from 55 to 64. That is, they varied from one-half to nearly one and one-half sigmas above the mean as given in the manual. Also, the means for both men and women are above 50 in all groups in the Hysteria component; in 13 out of 14 scores the means were above 50 in Paranoia; and in 11 out of 14 they were above in Hypomania. Most of the other components vary around 50 as the mean, except that in the Northwestern groups the scores are also significantly higher in Psychopathic Deviate and Schizophrenia components.

TABLE 2
THE PERCENTAGE OF NORTHWESTERN STUDENTS RECEIVING SCORES ABOVE 70

| Number of Categories | Per Cent | Cumulative Per Cent |
|---|---|---|
| 1 | 25 | 39 |
| 2 | 7 | 14 |
| 3 | 3 | 7 |
| 4 | 2 | 4 |
| 5 | 1 | 2 |
| 6 | 1 | 1 |

The extent and significance of this deviation from the norm is shown in Table 2, which gives the percentage of cases in which Northwestern scores were above 70. Of 600 cases, 233 (39 per cent) received a score in one or more categories of the test, which according to the manual indicates abnormality or at least is strongly suggestive of a psychopathic condition. A total of 14 per cent were above 70 in two categories, and 7 per cent in three categories. This is a very high percentage. There is no other supporting evidence to indicate anything approaching this high a degree of abnormality among Northwestern students. The data are not available to determine how the college groups from the other institutions would agree with the Northwestern data, but the mean and sigmas for the groups, where these data are available, indicate a somewhat smaller but nevertheless very high percentage of abnormality for the other groups.

Rather than accept these conclusions, it seems more reasonable to assume that the norms for certain categories of the test are too high, at least for college students. It is possible that college students are less inhibited and freer to give answers which indicate deviations from normality. This does not necessarily imply greater honesty, but they have less at stake than hospital patients, applicants for jobs, and applicants for entrance to college. Therefore, college students do not hesitate to check tendencies which others would guard against admitting even if true. But whatever the cause, these results indicate the extreme caution and the tentativeness that should be attached to any diagnosis based upon personality test scores alone.

### THE RELIABILITY OF THE MMPI

In addition to the problem of norms for a test, one of the other important considerations is the reliability of the test. The only report on the reliability so far as the writers can discover among the 234 titles in the "Bibliography on the MMPI" is to be found in the manual. It states: "The reliability coefficients of the scales so far developed in final form range between .71 and .83. These correlations were obtained by testing and retesting a group of 40 normals at intervals of a week or less up to several years." In order to make a further study of the reliability of the MMPI, the test was given in its two forms, individual and group, to 89 advanced students in psychology. The only difference between the two forms of the test is that in the group form, 16 questions are duplicated to facilitate machine

TABLE 3
RELIABILITY OF THE MMPI

| | *N* | *Hs* | *D* | *Hy* | *Pd* | *Mf* | *Pa* | *Pt* | *Sc* | *Ma* |
|---|---|---|---|---|---|---|---|---|---|---|
| Test-Retest | 89 | .29 | .81 | .39 | .79 | .71 | .67 | .70 | .55 | .56 |
| Split-Half | 97 | .78 | .58 | .47 | .46 | .73 | —.05 | .81 | .79 | .55 |

scoring. The other 550 items are identical in the two forms. To half the groups the group form of the test was given first, and to the other half the individual form was given first. The usual time between the two tests was one week; in a few cases it was as much as, but never more than, two weeks.

The test-retest and split-half reliability for the nine factors of the test are shown in Table 3. As will be noted from these results, two of the test-retest reliability coefficients are below .40, six are below .70, and one is .81. The mean reliability for the nine components of the test is .61. The corrected split-half reliabilities vary considerably from the test-retest reliabilities but have about the same mean value except for the Paranoia scale, which for some unknown reason is –.05. None of these are high coefficients of reliability and further suggest the caution that should be attached to any diagnosis based upon these tests alone. Only when other supporting evidence confirms the diagnosis can much weight be placed upon the results of the test. This is especially true for hypochondriasis, hysteria, paranoia, schizophrenia, and hypomania, five very common types of mental disorder.

### THE FORMS OF THE MMPI

The last part of this study is concerned with the amount of agreement in means for the individual and group forms of the MMPI. Only one study has compared the individual and group forms of this test. Weiner [4] gave tests, alternating the individual and group forms, to 200 consecutive veterans applying for vocational counseling at the Guidance Center in St. Paul. He found no significant difference between the two forms of the test when the raw scores were compared. As reported earlier in this present study, both the individual and group forms of the MMPI were given to 89 Northwestern University students. Table 4 shows that when the *K* correction is made and the standard scores are compared, only two of the nine scores approached significant differences. The group *Pd* score was higher, and the difference was significant at the 5 per cent level. The individual score for women was higher than the group form, and this difference was also significant at the 5 per cent level. These results indicate that there is little difference between scores on the individual and group forms of the test. In another study by one of the authors little difference in validity was found between the two forms of the test. While the two forms of the test do not correlate highly with each other, especially on certain categories of the

TABLE 4
COMPARISON OF MEANS FOR THE INDIVIDUAL AND GROUP FORMS OF THE MMPI

| | *N* | *Hd* | *D* | *Hy* | *Pd* | *Mf* | *Pa* | *Pt* | *Sc* | *Ma* |
|---|---|---|---|---|---|---|---|---|---|---|
| Weiner Scores | 100 | | | | | | | | | |
| Individual Test | | 51 | 51 | 55 | 53 | 53 | 50 | 50 | 51 | 57 |
| Group Test | 100 | 49 | 48 | 53 | 55 | 53 | 50 | 50 | 51 | 59 |
| Northwestern Scores | | | | | | | | | | |
| Individual Test | 89 | 50.2 | 50.6 | 54.8 | 53.5 | 59.8 Males<br>51.4 Females | 52.9 | 52.5 | 55.0 | 57.3 |
| Group Test | 89 | 51.5 | 51.5 | 55.9 | 56.6 | 63.5 Males<br>47.1 Females | 53.2 | 53.8 | 56.6 | 55.0 |
| *CR* | | 1.05 | 0.56 | 0.98 | 2.23 | 1.49 Males<br>2.37 Females | 0.25 | 0.97 | 1.06 | 0.75 |

test, there seems to be little clear evidence that one form has a higher validity than the other.

*Received February 13, 1951.*

### References

1. Blum, L. P. A comparative study of the students preparing for five selected professions including teaching. *J. exp. Educ.*, 1947, **16**, 31–65.
2. Brown, H. S. Similarities and differences in college populations on the Multiphasic. *J. appl. Psychol.*, 1948, **32**, 541–549.
3. Torrens, J. K. An investigation and evaluation of the Guilford Inventory of Factors STDCR with special reference to the MMPI. Unpublished paper, Univ. of Minnesota, 1944.
4. Weiner, D. N. Differences between the individual and group forms of the Minnesota Multiphasic Personality Inventory. *J. consult. Psychol.*, 1947, **11**, 104–106.

# AN ORIGINAL TEST OF SYMBOL ARRANGEMENT VALIDATED ON ORGANIC PSYCHOTICS

THEODORE C. KAHN
LOS ANGELES CITY SCHOOLS

THE idea of using plastic shapes of common objects for personality diagnosis occurred while observing purchases of such objects in a Los Angeles hobby supply shop. The plastic pieces were sold to be used as jewelry or decorations. An object's appeal depended on its size, color, material, and shape. Forms most liked by some people were ignored by others. The question arose—could the acceptance, rejection, and the manner of handling specific objects give insightful information regarding an individual's personality dynamics, his state of mental health, or his cerebral competence?

The materials of the Mosaic and World Tests as described by Bell [2, pp. 410–420, 468–473] had only superficial resemblance to the shapes in the hobby shop. The blocks used in the Goldstein-Sheerer [4] and the Hanfmann-Kasanin [5] tests are all geometric in design. The plastic forms used in the Doris Twitchell Allen Test [1] have ambiguous outlines with the exception of several geometric shapes. The literature on projective techniques did not reveal the existence of devices for personality diagnosis employing concrete and abstract, culturally structured, symbols such as hearts, crosses, anchors, butterflies—although art, archeology, and modern pictorial advertising are replete with them. On the other hand, the literature in psychology indicated that psychologists and psychiatrists have long been aware of the important role which symbols play in everyday living. Shaffer, writing on manipulative symbolism states: "all difficulties of life from the most trivial to the most grave adjustments, necessitate the use of these indicators in some form or other" [8, p. 74]. Disturbances in the ability to manipulate, recognize, and name object symbols are characteristic of many varieties of brain pathology. This is described by Nielsen [6, pp. 236–241] and acknowledged generally by neurologists.

In order to obtain a wide selection of shapes and materials and to include as large a geographic area as possible, hobby supply shops in Boston, New York, and Chicago were visited, until a sizable collection of objects for experimentation was accumulated. In the construction of the test, the number of items was finally narrowed down to fifteen. Versatility and variety of cultural associations were among the criteria used in the retention of specific items. Thus the horse shapes were eliminated, the dog shapes retained, since the latter offered a wider scope for acceptance or rejection. The faithfulness and loyalty of the dog is at one end—house pet, sled-puller, watchdog, and cattle-corraller is near the center—and being called the son of a female canine, undoubtedly, is at the other end of an acceptance-rejection continuum. Similarly, associations for the butterfly ranged from beauty to fickleness, and for the anchor, from travel and thus freedom of action to something which ties one down and thus destroys freedom of action. Tables 1 and 2 illustrate the popularity of the colors and shapes used in the Symbol Arrangement Test as judged by the frequency of occurrence in magazine advertisements and on the flags of the United Nations.

TABLE 1

NUMBER OF TIMES THE SHAPES ON THE KAHN TEST OF SYMBOL ARRANGEMENT APPEARED IN MAGAZINE ADVERTISEMENTS

| | |
|---|---|
| Number of different magazines reviewed | 56 |
| Number of advertisements studied | 5,726 |
| Times shapes used on test appeared in these advertisements | 18,828 |

TABLE 2

SIMILARITY BETWEEN THE OBJECTS USED IN THE KAHN TEST OF SYMBOL ARRANGEMENT AND THE FLAGS OF THE UNITED NATIONS

| | |
|---|---|
| Number of flags examined | 59 |
| Number of color similarities between test objects and flags | 166 |
| Number of similarities between test objects and designs on flags | 197 |

Another criterion in the selection of objects to be retained, was a similarity-with-difference factor to assist the subjects taking the test in the associative and projective processes. Three hearts, each differing in color, size, or material, are among the objects in the test. There are also two butterflies varying in shape, size, thickness, and color; three stars, two of which are identical; and three dogs, often seen as a family by the subject. Two dogs are alike in size and shape but not in color, and the third resembles one of the two in color but not in size or shape. If one wishes to combine the items for their cultural associations, one must sacrifice logical color combinations. If a color choice is made, the objects cannot be combined according to size or thickness.

This mutual exclusiveness of logical relatedness of the objects makes it possible for the examiner to ascertain the degree of correspondence between the subject's verbal intent and manipulation performance. Some of the shapes retained for the test approximate the popular responses of the Rorschach. Thus, the subject is able to manipulate winged objects, animals, and colors in plastic form detached from the amorphous matrix of an inkblot card. Only one of the fifteen objects in the Symbol Arrangement Test has an amorphous shape generally suggestive of the genitalia. Other representation retained in the test permit association ranging from things which occur in nature on several phylogenic levels to those which are man-made. The range in space extends from an anchor which may be located in the sea to outlines of the stars. There are also objects endowed with cultural meaning such as the cross and the hearts which transcend time and space.

An essential element in the Symbol Arrangement Test is a strip of brown felt containing fifteen consecutively numbered squares. This strip is intended to represent the part of the environmental background over which the subject has no control. The squares numbered from 1 to 15 conform to the consecutiveness imposed upon life by the immutable progression of time. The subject controls the manner of placing the plastic symbols in the squares, and the interaction of what he controls with what he does not, is believed to reflect in a miniature situation the subject's reaction to the conditions of life. The symbols which he places first or last, his reasoning in the arrangement of the objects, the meanings which he ascribes to them, and their juxtaposition as he places them on the strip, reveal information useful to the clinician.

The administration was standardized on 300 adults whose ages ranged from 18 to 87 years. It takes approximately thirty minutes to administer the test. Simplicity of administration, ease of reporting data, and objectivity of scoring were among the essential aims in the design.

TABLE 3

MATERIALS USED ON THE KAHN TEST OF SYMBOL ARRANGEMENT

Anchor—blue plastic (2 1/2″ × 1 1/2″ × 1/4″)
Butterfly—brown plastic (2 1/4″ × 1″ × 1/4″)
Butterfly—transparent plastic (2″ × 1 1/2″ × 1/32″)
Circle—transparent plastic (1 1/2″ diameter × 1/32″)
Cross—black plastic (1 1/2″ × 1 1/2″ × 1/4″)
Dog—black plastic (2″ × 1 5/8″ × 1/4″)
Dog—black plastic (1 3/4″ × 1 1/4″ × 3/16″)
Dog—white plastic (1 3/4″ × 1 1/4″ × 3/16″)
Heart—blue plastic (1 1/2″ × 1 1/4″ × 3/16″)
Heart—transparent plastic (2″ × 2″ × 1/32″)
Heart—red plastic (1″ × 1″ × 1/4″)
Parrot—green plastic (2 1/2″ × 5/8″ × 1/4″)
Star—transparent plastic (1 7/8″ × 1 7/8″ × 1/32″)
Star—red translucent plastic (1″ × 1″ × 3/16″)
Star—red translucent plastic (1″ × 1″ × 3/16″)
Brown felt strip—2″ × 30″. This strip is divided by white lines forming 15 2-inch squares, numbered consecutively 1–15 by white numerals 1/2″ in height.

## THE ADMINISTRATION OF THE TEST

Examiner: *In this box there are a number of plastic objects. What I am going to do is almost like playing a game.*

*I.* The examiner places the felt strip in front of the subject. Examiner turns the box containing the plastic objects upside down and lets the objects fall on the table.

Examiner: *Place these objects along this strip in any way.* (Examiner points to the objects and then to the strip. If subject requests further information, examiner replies: "In any manner.")

The examiner records the position of the objects as arranged on the felt strip, using abbreviations to denote the objects and numbers to designate the squares of the strip. With a small arrow he indicates the approximate degree of slant of those objects which were not placed horizontally. He notes the approximate time taken to complete the task, and the direction from which the subject proceeded in arranging the objects (i.e., left to right, right to left, or mixed).

Examiner: *Why did you line then up this way?* Examiner records the reply. He continues: *That's fine. Now name each object from one to fifteen.* If the subject requests further information, the examiner says: "Tell me what each object is." Examiner records these responses and then picks up the strip, letting the objects fall on the table. He gathers the objects together and again stretches the strip out in front of the subject.

*II.* Examiner: *Now line up the objects once more in any manner.* (Examiner explains if necessary, "In any manner you wish.")

Examiner records the subject's performance as in *I.*

Examiner: *Why did you line them up this way this time?* The subject's response is recorded and the examiner continues with: *A flag can represent a country. A horseshoe can stand for good luck. A torch can symbolize knowledge or freedom. Tell me what each of these objects could stand for, represent, or symbolize. Go from one to fifteen.* (In the case of children and mental defectives it has been found necessary to add: "What do these objects make you think of? What comes to your mind when you look at them?" If the subject cannot symbolize or abstract, free association or naming is accepted. The subject is encouraged as he proceeds. If the subject is hesitant regarding his abstractions, the examiner may nod or say: "That's fine.") The examiner records the responses, gathers up the strip with the objects, piles them together on the table, and then stretches out the strip again in front of the subject.

*III.* Examiner: *Now line up the objects exactly as you had them just before. See how close you can come to it.* (If there is a question, the examiner says: "See whether you can remember how you had them lined up the second time you did it.")

The examiner records the subject's performance as in *I* and may say: "You got all of them right," or "You got some of them right," or give some other evidence of encouragement.

The examiner removes the transparent heart, the transparent star, the transparent butterfly, and the transparent circle from the strip and holds them. He hands the transparent heart to the subject, and says: *Place this over any one of the objects remaining on the strip.* The examiner notes the response and then removes the transparent heart from the object on which it has been placed and retains it. He continues, using the same procedure for the transparent star, the transparent butterfly, and then the transparent circle, giving the same instructions in each case. He then gathers up the strip with the objects, piles them together on the table, and adds the transparent pieces he has held. The examiner again stretches the strip out in front of the subject.

*IV.* Examiner: *This time I want you to arrange the objects on the strip according to how each one appeals to you. Place the object you like most in the first square. Put the one you like next best in the second square. In the third square, place the object you like third best. Continue in this manner along the strip, placing the objects you like least at the end.* (Examiner points to the squares on the strip in order to help the subject grasp the idea. He records the subject's performance as in *I.*)

Examiner: *Why did you place this one first?* Examiner points to the first square. (If necessary, he inquires: "Why did you like this best of all?") He records the response.

Examiner: *Why did you place this next?* (He points to the second one or, "Why did you like this one next best?") The response is recorded.

Examiner: *Why did you like this third best?*

Examiner records response. (The subject is asked to express himself regarding three different concepts. For example, if he puts all three hearts at the beginning, saying that they all stand for love, the examiner questions the subject further until the latter has given three different reasons for liking objects.) The examiner then points to the object on square 15.

Examiner: *Why did you like this least of all?*

Examiner records the response and continues until three reasons for disliking objects are given. Some subjects cannot or will not give any reasons for disliking objects. In that case the examiner records this as a response.

*V.* Examiner: *This is the last time I'll ask you to do it. Once more I am going to ask you to line up the objects in any way.* The performance is recorded.

Examiner: *Why did you line them up this way this time?* The response is recorded, and the administration of the test has been completed.

*Testing the limits.* The examiner should test the limits if the subject did not arrange the objects according to some logical plan but consistently re-

plied that he placed them at random or had no reason for placing the objects as he did. In that case the examiner may ask: "Could you arrange the objects according to some method or idea?" If the subject is still confused, the examiner may add: "Can you put together all the objects that are related in any way?" Psychotics have shown difficulty in seeing the relationship of the objects and often were incapable of making logical arrangements based on similarities of shapes, color, design, or abstraction. This is consistent with psychotic performance on other tests of concept formation such as the Goldstein-Sheerer battery [4].

The examiner may observe that the subject has placed all the dogs together but has separated the hearts, stars, or butterflies. When some of the objects of similar content have been placed together but others have not, the examiner may ask for reasons. Caution must be used as this kind of questioning is threatening to some subjects.

In neurological examinations the examiner may ask the subject to place all those objects together which are alike in some way. Since the similarity of the objects to each other is graduated from identity to only abstract likeness, a rough scale of capacity to recognize relationships is available. To assist in the identification of sensory-motor impairment, the examiner may give the objects to the subject under a table and ask the latter to draw or name them. Aphasics are able to point to objects named by the examiner but are incapable of naming them when the examiner points them out.

### SCORING CRITERIA

Scoring criteria consist of common motor and verbal responses in the performance of the test which lend themselves to ready observation and definition. Some of these are completely objective. They are: the three shapes most liked and the three least liked, the number of similarly shaped objects placed together, the number of objects placed in slanting or upside-down positions on the strip, the number of errors in recall, the number of divergencies in the naming responses from a given norm, the time required to complete arrangement, the direction in which the objects are arranged (left to right, right to left, or mixed), the shapes over which the transparent objects are placed, the number of times objects overlap or are partly off the strip, and the frequency with which the subject touches or points to the object during the naming and symbolizing.

The remaining scoring criteria may be considered as "semi-objective." The scoring here follows the Wechsler-Bellevue type of evaluation by matching responses with specific criteria weighted according to normative data. Included in these are level of symbolization, reasons given by the subject for arranging the objects as he did, and reasons given for liking or disliking specific objects. Table 4 illustrates the nature of the evaluation of the symbolization responses.

TABLE 4

EXAMPLES OF LEVELS OF SYMBOLIZATION FOR THE ANCHOR OBJECT

| Symbolization | Type of Response |
|---|---|
| 1. "A baby" | Bizarre |
| 2. "Don't know," "Can't do it" | No symbolization |
| 3. "An anchor" | Naming |
| 4. "To keep a ship from drifting" | Description of function |
| 5. "Navy insignia" | Association with fidelity to form |
| 6. "The seven seas," "ships in the harbor" | High abstraction, tangible level |
| 7. "Transportation," "security" | High abstraction, intangible level |

### VALIDATION ON A BRAIN-DAMAGED PSYCHOTIC GROUP

In order to determine how effectively the Symbol Arrangement Test could differentiate between organic psychotics and nonpsychotics, a study was undertaken at the Brentwood Veterans Administration Hospital in West Los Angeles [6]. Fifty male patients were selected whose psychiatric diagnosis of brain damage was supported by encephalographic records and case histories. Fifty male nonpsychotic veterans were matched with the hospital group for age, race, religion, and nationality. Criteria for nonpsychosis were cutoff scores on the Minnesota Multiphasic Personality Inventory, personal histories, and psychological interviews. Criteria for non–brain-damage included medical histories free from brain damage, supported by the opinion of a physician or psychiatrist. The two groups were compared as to occupational background, IQ scores, marital status, and educational level.

TABLE 5

SIGNIFICANCE OF VARIABLES IN A COMPARISON OF BRAIN-DAMAGED PSYCHOTICS WITH NONPSYCHOTICS ON THE KAHN TEST OF SYMBOL ARRANGEMENT

| Category of Variable | Section of Administration | Not Significant | No. of Responses Significant between the .05 and .10 Levels | No. of Responses Significant at < .05 Level | No. of Responses Significant at < .01 Level | No. of Responses within Category |
|---|---|---|---|---|---|---|
| 1. Deviation from norms in the naming of objects | I | – | – | – | 15 | 15 |
| 2. Levels of symbolization | II | 2 | – | – | 5 | 7 |
| 3. Similarly shaped objects placed in juxtaposition | IV | – | – | – | 5 | 5 |
| 4. Reasons for liking objects | IV | 2 | 2 | 1 | 5 | 10 |
| 5. Reasons for disliking objects | IV | 2 | 1 | 2 | 5 | 10 |
| 6. Lifting or touching objects in naming and symbolization | I & II | 4 | 1 | – | – | 5 |
| 7. Fragment of circle placed over other object (not mentioned in this article) | II | 11 | 2 | 3 | – | 16 |
| 8. Transparent heart placed over other object | III | 9 | 1 | 1 | 1 | 12 |
| 9. Transparent star over other object | III | 9 | – | 2 | 1 | 12 |
| 10. Transparent butterfly over other object | III | 8 | 2 | – | 2 | 12 |
| 11. Transparent circle over other object | III | 8 | 1 | 2 | 1 | 12 |
| 12. Objects liked most | IV | 11 | 2 | 2 | – | 15 |
| 13. Objects liked least | IV | 10 | 2 | 2 | 1 | 15 |
| 14. Reasons for arranging objects | I, II, & V | 5 | – | 2 | 6 | 13 |
| 15. Number of errors in recall | III | – | – | – | 1 | 1 |
| 16. Amount of time elapsed before naming and symbolization | I & II | – | – | 1 | – | 1 |
| 17. Time required to arrange objects | I–V | 14 | – | 3 | 3 | 20 |
| 18. Directions from which the objects were arranged | I–V | 3 | 3 | 2 | – | 8 |
| 19. Placing objects in slanting or upside down position on strip | I–V | – | – | – | 8 | 8 |
| Total | | 98 | 17 | 23 | 59 | 197 |

One hundred and ninety-seven responses were compared by *t*-ratios. Ninty-nine of these were found to be significant, 59 at better than the 1 per cent level of confidence. Eighteen responses yielded *t*-ratios of over 5.0. Significant differentiating performance by the psychotics included a larger proportion of the following: placing objects in slanted and upside-down positions on the strip, objects overlapping, dissimilarly shaped objects placed together, errors in recall, time taken in performance, direction of arrangement of objects, and level of symbolization and reasoning.

Scores were derived from *t*-ratios computed from the difference between the means and the proportion of responses of the two groups. The mean total score for the psychotic group was —221 with a range of +28 to —440. The

mean total score of the nonpsychotics was +104 with a range of —86 to +240. In only 5 of the 100 cases was there a slight overlapping of scores. A statistical comparison of total scores for the two groups yielded a $t$ of 17.6.

## RELIABILITY AND OTHER DATA

A comparison of 25 total scores on the Symbol Arrangement Test with IQ scores of the same group showed no significant correlation. A test-retest method for ascertaining reliability yielded an $r$ of .95 on 25 cases of an unselected male group tested after a lapse of six months. Ten records were scored separately by two examiners, and the resulting correlation of total scores yielded an $r$ of .97.

Results of another study using the Symbol Arrangement Test with 50 schizophrenics [3] indicates that the technique can differentiate between organic psychotics, schizophrenics, and nonpsychotics.

Among the differences noted in the performance of the organic and schizophrenic groups was the greater tendency of the former to put dissimilar objects together, to slant objects in placement, to show deviations or inability in the naming of the objects, to use lower levels of abstractions, to crowd objects together, and to have objects overlap or extend beyond the edges of the strip. The schizophrenics gave a larger number of emotionalized and pure color responses in the symbolization. They touched and handled the objects more frequently than did the organic group, and showed greater ability in recognition of shape similarities.

## SUMMARY

A new test of Symbol Arrangement has been presented. It was designed to assist clinicians in the exploration of personality dynamics and to yield objective criteria for differential diagnosis. The test consists of fifteen plastic shapes of various colors and with a variety of cultural associations. The administration takes approximately thirty minutes. It requires the subject to arrange the pieces on a felt strip several times under varying conditions.

The technique has been found suitable for use with organic psychotics in a neuropsychiatric hospital. It differentiated statistically between the hospital group of 50 male veterans diagnosed as psychotics with brain damage and a matched group of nonpsychotics. Scores were based on weights derived from $t$-ratios. The study indicates that the Symbol Arrangement technique may be of value in the identification of organic psychosis.

*Received June 8, 1951.*
*Early publication.*

## REFERENCES

1. ALLEN, D. T. *Three-dimensional apperception test.* New York: Psychological Corp., 1948.
2. BELL, J. E. *Projective techniques.* New York: Longmans, Green, 1948.
3. FILS, D. H. Comparative performance of schizophrenics and normals on a new projective test of object-symbol arrangement. Paper read at the Western Psychol. Ass., San Jose, Calif., April 1951.
4. GOLDSTEIN, K., AND SHEERER, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, **53**, No. 239.
5. HAUFMANN, EUGENIA, AND KASANIN, J. Conceptual thinking in schizophrenia. *Nerv. ment. Dis. Monogr.*, 1942, No. 67.
6. KAHN, T. C. Comparative performance of psychotics with brain damage and nonpsychotics on an original test of symbol arrangement. Paper read at the Western Psychol. Ass., San Jose, Calif., April 1951.
7. NIELSEN, J. M. *Clinical neurology.* New York: Hoeber, 1947.
8. SHAFFER, L. F. *The psychology of adjustment.* Boston: Houghton Mifflin, 1936.

# NEW BOOKS AND TESTS

## Books

BLAKE, ROBERT R., AND RAMSEY, GLENN V. (Eds.) *Perception: An approach to personality.* New York: Ronald, 1951. Pp. viii + 442. $6.00.

This volume is based on thirteen papers delivered at a clinical psychology symposium at the University of Texas during 1949–50. Defining perception broadly, Blake and Ramsey asked each of twelve persons to present his research findings and thinking in the general areas of personality and interpersonal relations. Individual chapters were contributed by Beach, Bronfenbrenner, Bruner, Cameron, Dennis, Frenkel-Brunswik, Hilgard, Klein, Korzybski, Miller, Morgan, and Rogers. The topics thus included range from the physical and chemical determinants of behavior, through developmental and social factors in perception, to unconscious processes, behavioral pathology, and psychotherapy. Although the quality of the individual chapters varies, the editors managed to achieve considerable integration of this range of materials. This book will appeal to a wide audience and should stimulate further integrative efforts in research and in theory development.—*E. L. K.*

BRUN, RUDOLF. *General theory of neuroses.* (Trans. by Bernard Miall). New York: International Universities Press, 1951. Pp. vii + 469. $10.00.

Thirty-two lectures on the neuroses, by an outstanding Swiss psychiatrist, are here translated into English for the first time. In the main, Brun follows Freud, although there are a few conspicuous departures. With Freud, he holds that the "actual" neuroses—neurasthenia and anxiety neurosis—have an organic basis. He rejects Freud's "sexual excesses" theory of such neuroses, however, and ascribes them to pathological sensitization processes in the vegetative nervous system. In keeping with this theory, the recommended treatment is medical rather than psychotherapeutic. Brun's theory of instincts expands Freud somewhat; his discussions of psychoneuroses and of the mechanisms of symptom formation do not differ significantly from Freud's. Although much of Brun is already quite familiar to anyone who has read widely in analytic literature, there are more than a few distinctive and valuable passages. On the whole, the book is worth reading.

COLBY, KENNETH MARK. *A primer for psychotherapists.* New York: Ronald, 1951, Pp. xii + 167. $3.00.

This little book is intended for beginners in psychotherapy, among them "internes or residents in psychiatric hospitals and clinics, practicing psychiatrists, clinical psychologists, and psychiatric social workers." At least, the author's preface says so; but, some twenty pages later, all this is forgotten, and medical training is assumed for all psychotherapists. Not many pages later, psychiatrists are warned not to administer physical examinations because "By the time one has become something of a psychotherapist, his medical diagnostic judgment has suffered a disuse atrophy of such a proportion that he is really no longer a reliable medical man." Make up your mind, Doctor! The volume has many snatches of therapeutic interviews, which are inherently interesting and which illustrates some of the author's techniques in treatment. It is difficult to determine for whom this book is useful. It is too simple for therapists and too complex for beginners. It presupposes a rather thorough knowledge of Freudian psychology, and one with that knowledge will not benefit too much from simple expositions of methods of beginning and ending therapy.—*M. K.*

CROW, LESTER D., AND CROW, ALICE. *Mental hygiene.* (2nd Ed.) New York: McGraw-Hill, 1951. Pp. xviii + 433. $4.50.

As a text in mental hygiene, this book has few merits and many shortcomings. It lacks any systematic theoretical orientation, but depends on an unintegrated miscellany of borrowings. There is little evidence of scholarship. The few studies drawn upon are quite likely to be cited from secondary sources, when the originals are easily available. A ponderous style of writing makes easy concepts seem difficult. Finally, in spite of the asserted emphasis on positive mental hygiene, there is grave doubt that the book is even hygienic. A few exam-

NOTE: Some reviews in this issue were prepared by the Associate Editors, who may be identified by their initials. Unsigned reviews are by the Editor.—*L. F. S.*

ples suffice. An infant's needs are "mainly physical" (p. 73). If a six-year-old still wets the bed, "there is probably a physical difficulty that needs attention" (p. 77). In descriptions of the symptoms of mental illness, phobias are sandwiched between hallucinations and delusions (p. 372), yielding an implication likely to terrify the numerous readers who have irrational fears. It is to be hoped that students will be spared contact with this potentially harmful book.

DUNSMOOR, CLARENCE C., AND DAVIS, OLIVER C. *How to choose that college.* Boston: Bellman Publishing Co., 1951. Pp. 52. 90¢.

This attractive and informative manual, designed to help students choose a college and secure admission to it, can be used individually or in high school guidance classes.

FINEGOLD, S. NORMAN. *Scholarships, fellowships, and loans.* Vol. II. Boston: Bellman Publishing Co., 1951. Pp. 312. $5.00.

Supplementing and not replacing the first volume published in 1949, this compilation gives full information concerning 245 agencies that grant scholarship aid for college, vocational, and graduate study. An introductory section on career planning, and three indexes, help to make the book serve its users. It will be an invaluable aid in educational guidance programs.

FISCHER, SIEGFRIED. *Principles of general psychopathology.* New York: Philosophical Library, 1950. Pp. xxi + 327. $4.75.

In this small volume, the author, formerly Professor of Psychiatry and Neurology at the University of Breslau, attempts to present the fundamentals of psychopathology and to integrate them by means of a dynamic theory of psychology which is at the same time critical of psychoanalysis. Psychopathology is defined as the science of "the regularities of pathological mental happenings." According to the author, there are only two sources of knowledge of mental happenings: introspection and empathetic observation. Approximately half of the forty references are to European authors including Ebbinghaus and Lipps.—*E. L. K.*

FREEMAN, LUCY. *Fight against fears.* New York: Crown Publishers, 1951. Pp. 332. $3.00.

Miss Freeman, who is well known to psychologists for her excellent reporting of psychological topics in the *New York Times,* here gives a highly personal account of her own psychoanalysis. It is a readable human document that conveys a real sense of how it feels to fear and then to overcome fears. It also reveals a tolerant, anonymous analyst who does a good job without the aid of technical jargon.

GINZBERG, ELI, GINSBERG, SOL W., AXELRAD, SIDNEY, AND HERMA, JOHN L. *Occupational choice.* New York: Columbia Univ. Press, 1951. Pp. ix + 271. $3.75.

A preliminary report of a significant research on vocational choice, using analyses of extensive recorded interviews. So far, the groups interviewed do not go beyond the age of college students, but include high and low socioeconomic levels. The data point to a general theory that occupational choice is a developmental process, involving not one decision but a series of decisions, the effects of which may be largely irreversible. Typical stages in the development of vocational choice have been identified, and are related to age, education, ability, and sex. Although the conclusions are not startling, they are sound and serve to enlarge our understanding of one of life's most important sequences. Further reports of the studies of Dr. Ginzberg's group will be awaited with interest.

GUETZKOW, HAROLD. (Ed.) *Groups, leadership and men.* Pittsburgh: Carnegie Institute Press, 1951. Pp. ix + 293. $5.00.

A report on researches sponsored by the Human Relations and Morale Branch of the Office of Naval Research, 1945–1950, as summarized and discussed at a conference held in 1950. The reports are grouped in three sections. The first section deals with research on group behavior, including problems of group morale, and the effect of human relationships on the loyalties of the worker to management and to the union. The second section contains research on leadership and the behavior of leaders. The third section deals with individual behavior. Reports are given on new devices for the measurement of achievement, on motives, and on an evaluation of neuropsychiatric screening. It is a difficult book to evaluate. An attempt has been made to give the impact of the various papers on the members of the conference, and the reader has the feeling of being present and listening to the discussions. There is much variation in the quality and potential significance of the papers, but they are stimulating and often challenging. No final solutions are reached. In fact, most of the papers might be termed "progress reports." It is a book that should be read by all who are planning research or study in the field of human relations.—*B. M. L.*

KITSON, HARRY D. *How to use your mind.* (4th Ed.) Philadelphia: Lippincott, 1951. Pp. 196. $2.50.

A fourth edition of a book that, since 1926, has offered realistic and honest applications of psychology to personal efficiency. There are new chapters on effective reading, increasing one's vocabulary, taking psychological tests, and mental hygiene.

LOVELL, HAROLD W. *Hope and help for the alcoholic.* Garden City, N. Y.: Doubleday, 1951. Pp. 218. $2.75.

An excellent nontechnical book about alcoholism and its treatment, which can be recommended to alcoholics and to their families.

MAAS, HENRY S. (Ed.) *Adventure in mental health.* New York: Columbia Univ. Press, 1951. Pp. xiii + 334. $4.50.

This volume reports the psychiatric social work in various branches of military service. Sixteen workers collaborate to describe their problems and practice in these settings, and the implications of this experience for current civilian practice. Among the chapters are those entitled: combat division, convalescent hospital, disciplinary barracks, professional association in a world crisis, and selective service medical survey program. The collection is a good report of the present-day social worker on the psychiatric team. The infrequent reference to the psychologist in the book is more evidence for the need for both professions to know each other better.—*F. McK.*

MONS, W. *Principles and practice of the Rorschach personality test.* (2nd Ed.) Philadelphia: Lippincott, 1951. Pp. 176. $4.00.

The second edition of this compact and clear British text differs little from the first edition (see *J. consult. Psychol.*, 1949, **13**, 309), except that civilian cases have been substituted for the military examples.

MUNN, NORMAN L. *Psychology.* (2nd Ed.) Boston: Houghton Mifflin, 1951. Pp. xvi + 624. $4.50.

Munn's widely used text has been revised thoroughly. There is somewhat less material on animal experiments and on the nervous system, and more on human experiments and applications. The chapter on acquired motivation is especially well handled in the revision. The book remains a solid instructional resource, equally useful for general education and for a first study of psychology by students who will continue professional preparation.

[KOHN, NATHAN, JR., CHM.] *1951 directory of vocational counseling services.* St. Louis Mo.: National Vocational Guidance Association, Ethical Practice Com., Box 64, Washington Univ., 1951. Pp. 125. $1.00.

The second issue of the list includes 166 agencies, services, and individuals that have been examined and approved. The NVGA is to be congratulated on its continued sponsorship of this project of mature self-evaluation.

RAPAPORT, DAVID. (Ed.) *Organization and pathology of thought.* New York: Columbia Univ. Press, 1951. Pp. xviii + 786. $10.00.

Rapaport has gathered together and translated twenty-seven European contributions on the organization and pathology of thought. The original authors include such names as Claparède, Lewin, Piaget, Freud, Bleuler, Schilder, and Kris. Each selection is accompanied by copious interpretative notes by the editor. He intentionally has selected writings not easily accessible in English and writings that go beyond the current American tendency to limit the study of the thought processes to concept formation and the conventional learning theory approach. In a personal selection such as this there are bound to be some inclusions and omissions as well as some editorial interpretative comment that the individual reader may question. The book remains, however, a valuable and stimulating addition to our literature on the subject. There is a brief introductory section in which Rapaport explains and justifies his purpose and sets his material in context. A longer concluding section in which the editor attempts some theoretical integration is, in this reviewer's opinion, not quite as successful as the rest of the book. The total result, while perhaps not necessarily required reading, is certainly to be recommended as stimulating, challenging, and worth while.—*W. A. H.*

REITMAN, FRANCIS. *Psychotic art.* New York: International Universities Press, 1951. Pp. x + 180. $4.50.

*Psychotic Art* is a compact, well-written treatment not only of psychotic art, but of the dynamics of art production in general. It is scholarly, interesting, and well documented. The author, however, seems as much concerned with the annihilation of several schools of psychology as with discussion of the art production of the mentally ill. Freudian, Jungian, and typological psychology come in for violent attack. Reitman uses this volume as a vehicle for presenting his position that psychosis, particularly schizophrenia, is organically based. In spite of the physiological bias, this study contains a wealth of material on artistic manifestations of personality deviation. Adherents of other schools of psychology, however, will draw quite different conclusions from the same material.—*M. K.*

SCHNEIDERS, ALEXANDER A. *Introductory psychology.* New York: Rinehart, 1951. Pp. xv + 461. $4.00. *Workbook in introductory psychology.* New York: Rinehart, 1951. Pp. 74. $1.25.

A brief text in general psychology, intended for the one-semester introductory course. The author's intention to center the text's organization on personality and adjustment is laudable, but most of the subject matter is conventional. The workbook

provides objective test items, questions, and projects on each chapter.

SECHEHAYE, M. A. *Symbolic realization.* (Trans. by B. and H. Wursten). New York: International Universities Press, 1951. Pp. 184. $3.25.

A vivid account of the successful treatment of a case of schizophrenia by a Swiss woman analyst. The method of "symbolic realization" relies on communicating with the patient in terms of her own symbols, when she cannot use the conventional language symbols of normal people. The book makes a rich contribution to the understanding both of schizophrenia and of psychotherapy.

SMITH, GLENN E. *Principles and practices of the guidance program.* New York: Macmillan, 1951. Pp. xi + 379. $3.25.

A soundly organized and well-written text about guidance programs and services in schools. The author views the guidance counselor as a leader who sensitizes the educational community to guidance needs and organizes services, rather than solely as a worker with individual pupils. Chapters survey the history of guidance, and the main aspects of programs—inventories, information, surveys, counseling, placement, follow-up, and evaluation. The book gives the reader a good picture of the better current guidance practices at the high school level, and is a suitable text for a survey course.

STEVENS, S. S. (Ed.) *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. xi + 1436. $15.00.

The long-awaited *Handbook* fulfills the expectations of psychologists. After seventeen years, we again have a comprehensive survey of psychology, the product of thirty-four well-qualified authors. The coherence of style and the readable presentation of the intrinsically difficult materials show that the editor exercised his functions with care and skill. What service does the *Handbook* offer to clinical psychologists? First, quite naturally, the clinician will look for sections most directly applicable to his interests and tasks. He will find very explicit use for the chapters on homeostasis (Dempsey), development (Carmichael), genetics (Hall), motivation (Beach, N. E. Miller), and emotion (Lindsley). Only a little peripheral to clinical needs are the sections on learning (Hilgard, Brogden, Hovland, Spence, Morgan), cognitive processes (Leeper), and speech and language (G. A. Miller). The four last chapters of the *Handbook* are on areas of applied experimental psychology, particularly as used in industry and in the armed forces—selection (Bechtoldt), training (Wolfle), equipment design (Fitts), and work (R. H. Seashore). In view of the inclusion of these "applied" chapters, clinical and social psychologists may wonder why some topics were omitted. To name only a few examples, studies of the influences of motivation and attitude upon perception, of the development of personality, and of conflict and frustration, are sufficiently well developed to warrant summaries, and are as "general" in their appeal as some topics included. The omission of such areas, of course, can be attributed to the pressure for space. A handbook cannot be entirely encyclopedic, and the editor had to draw the line somewhere. Clinical psychologists who want to be more than specialized technicians need this volume as much as do experimental psychologists. It is a rich summary of many of the most important foundations of clinical research and practice.

STRANG, RUTH. *An introduction to child study.* (3rd Ed.) New York: Macmillan, 1951. Pp. xi + 705. $4.75.

In its third edition, this widely used book continues to present a clear and comprehensive account of child and adolescent development. The book is "easy reading," as it should be for its intended audience, but its clarity and literary grace do not obscure its penetrating scholarship. It also differs from "popular" books on childhood in its lack of prescriptive instructions for child rearing. The author trusts the reader to arrive at sound conclusions when he has a basis for real understanding. The bibliographies are extensive and new, and there are lists of films and other training aids.

ZANGWILL, O. L. *An introduction to modern psychology.* New York: Philosophical Library, 1950. Pp. xi + 227. $3.75.

A pocket-sized survey of the whole field of psychology, originally published in Great Britain as one of a series of "home study books." Opening with a historical chapter, the book proceeds to brief but penetrating surveys of experimental psychology, physiological psychology, mental measurement, psychoanalysis, and personality. The author does not balk at complicated topics but succeeds in making them clear and interesting. The book stands in refreshing contrast to what usually passes for "popular" literature in psychology.

## TESTS

BENNETT, GEORGE K. *Mechanical Comprehension Test, Manual* (Rev. Ed.) New York: Psychological Corp., 1951. Pp. 8.

The revised manual contains added data on the correlation of the test with various criteria, and new norms based on large and appropriate groups of students and employees.

BENNETT, GEORGE K., SEASHORE, HAROLD G., AND WESMAN, ALEXANDER G. *Validation of the Differential Aptitude Tests, fourth research report.* New York: Psychological Corp., 1951. Pp. 8.

The fourth research report includes four tables of correlations of test scores with school grades, and the report of a special study following a group of pupils over a period of two and one-half years. The relatively good predictive value of the tests over that interval is of interest.

HATHAWAY, S. R., AND MCKINLEY, J. C. *Minnesota Multiphasic Personality Inventory, Manual.* (1951 Revision). New York: Psychological Corp., 1951. Pp. 31. $1.50.

The revised *Manual* is a compact guide for the use of the MMPI, including the developments of the past five years. There are directions for the use of both the individual (card) form and the group (booklet) form. New features include instructions for the use of the *K* scale and of the *Si* (social introversion-extraversion) scale, and a brief account of profile coding. Although there is a short discussion of the interpretation of profiles, no attempt is made to summarize the extensive clinical and research literature on the MMPI. The selected bibliography contains 47 references, mainly recent ones.

KUDER, G. FREDERIC. *Examiner manual for the Kuder Preference Record—Vocational,* Form C. (2nd Revision). Chicago: Science Research Associates, 1951. Pp. 32. 25¢.

The current revision of the manual for the Kuder (see *J. consult. Psychol.,* 1949, **13**, 67) contains new norms for boys, girls, and men based on larger samples, added job profiles, and a revised table of occupations classified according to major interests.

KUDER, G. FREDERIC, AND CRAWFORD, LURA E. *Kuder Book List.* Grades 7–12. 1 form. Book list ($2.35 per 25), with instructor's guide, pp. 4. Chicago: Science Research Associates, 1951.

The *Book List* is not a test, but an instructional device for personal development to be used with the *Kuder Preference Record—Vocational.* It consists of over five hundred book titles, classified under the ten interest areas defined by the Kuder blank. From his Kuder profile, the youngster is led to a list of books that probably will interest him. Within each interest area, the books are divided into four levels of difficulty, corresponding approximately to fifth-, seventh-, ninth-, and eleventh-grade reading proficiency. The purpose of the list is to develop enthusiasm for reading among students by introducing them to books that they will enjoy.

## PSYCHOLOGICAL MONOGRAPHS: GENERAL AND APPLIED

### Volume 64, 1950

**Patterns of Personality Rigidity and Some of Their Determinants.** Seymour Fisher Elgin State Hospital. #307, $1.00

**The Value of an Oral Reading Test for Diagnosis of the Reading Difficulties of College Freshmen of Low Academic Performance.** Charles A. Wells, American International College. #308, $1.00

**Rorschach Responses Related to Vocational Interests and Job Satisfaction.** Solis L. Kates, Michigan State College. #309, $1.00

**Symbol Elaboration Test (S.E.T.): The Reliability and Validity of a New Projective Technique.** Johanna Krout, Chicago Psychological Institute. #310, $2.00

**Changes in Responses to the Minnesota Multiphasic Inventory Following Certain Therapies.** William Schofield, University of Minnesota. #311, $1.00

**A Scale for Measuring Teacher-Pupil Attitudes and Teacher-Pupil Rapport.** Carroll H. Leeds, Furman University. #312, $1.00

**The Nature and Efficacy of Methods of Attack on Reasoning Problems.** Benjamin Burack, Roosevelt College. #313, $1.00

**The Validity of a Multiple-Choice Projective Test in Psychopathological Screening.** Martin Singer, VA Hospital, Long Island. #314, $1.00

**A Normative Study of the Thematic Apperception Test.** Leonard D. Eron, Yale University. #315, $1.50

**Experimentally Induced Variations in Rorschach Performance.** Edith E. Lord, Arizona State Department of Health. #316, $1.00

**An Evaluation of Personality-Trait Ratings Obtained by Unstructured Assessment Interviews.** Ernest C. Tupes, United States Air Force. #317, $1.00

The 1950 volume of the Psychological Monographs consists of eleven separate issues. Orders for any of these can be placed separately at the prices listed above, or the entire volume can be ordered for $6.00.

## American Psychological Association

**1515 Massachusetts Avenue N.W., Washington 5, D.C.**

# JOURNAL OF CONSULTING PSYCHOLOGY
# AVAILABLE BACK ISSUES

The issues listed below are available for sale

| YEAR | VOLUME | JAN FEB | MAR APR | MAY JUN | JUL AUG | SEP OCT | NOV DEC | PRICE PER NUMBER | PRICE PER VOLUME |
|---|---|---|---|---|---|---|---|---|---|
| 1937 | 1 | 1 | 2 | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1938 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1939 | 3 | 1 | 2 | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1940 | 4 | 1 | 2 | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1941 | 5 | 1 | 2 | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1942 | 6 | 1 | – | 3 | – | – | 6 | $.60 | $1.80 |
| 1943 | 7 | – | 2 | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1944 | 8 | 1 | – | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1945 | 9 | 1 | 2 | 3 | 4 | 5 | 6 | $.60 | $3.00 |
| 1946 | 10 | 1 | 2 | 3 | 4 | 5 | – | $.60 | $3.00 |
| 1947 | 11 | – | – | 3 | 4 | 5 | 6 | $.60 | $2.40 |
| 1948 | 12 | 1 | 2 | 3 | 4 | 5 | 6 | $1.00 | $5.00 |
| 1949 | 13 | 1 | 2 | 3 | 4 | 5 | 6 | $1.00 | $5.00 |
| 1950 | 14 | 1 | 2 | 3 | 4 | 5 | 6 | $1.00 | $5.00 |
| 1951 | 15 | By Subscription, $5.00 | | | | | | | |

Price for Volumes 1 through 14 $46.20

Table based on inventory of January 1, 1951

The Journal of Consulting Psychology is of particular interest to clinical psychologists, psychiatrists, school psychologists, and persons who are engaged in counseling and guidance work.

The price per volume of the Journal of Consulting Psychology was $3.00 between 1937 and 1947. It is now $5.00 per volume and $1.00 per issue. For foreign orders, $.25 per volume should be added. Address orders to:

AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 Massachusetts Ave. N. W.
Washington 5, D. C.

# NEUROLOGY & PSYCHIATRY

**Section VIII of Excerpta Medica**

The Neurology Section includes abstracts of all articles on neurological surgery in addition to general clinical neuro-physiology.

The section on Psychiatry covers the topics usually contained within the meaning of the word. In addition, the following special branches are dealt with fully:

* **Psychoanalysis, individual and analytical psychology**
* **Sexology, criminology, alcoholism, and drug addiction as related to psychiatry**
* **Social and industrial psychology and psychiatry, vocational guidance and mental testing**
* **Psychometrics**
* **Personality testing and the Rorschach test**
* **Heredity and statistical studies as they apply to psychiatry**
* **Mental defect and epilepsy**

Annual volume of 1000 pages, including an index classified by author and subject.

**Subscription price $22.50 per annual volume**

**THE WILLIAMS & WILKINS COMPANY**

Agents for the United States, Canada and Central America

**Mt. Royal and Guilford Aves. Baltimore 2, Maryland**

---

Please enter my subscription to Section VIII, NEUROLOGY AND PSYCHIATRY, OF EXCERPTA MEDICA. I enclose $22.50.

Name.....

Address.....

.....

THE PSYCHOLOGICAL CORPORATION

**New . . for counselors . . .**

*A Casebook for the Differential Aptitude Tests*

# Counseling from Profiles

GEORGE K. BENNETT : HAROLD G. SEASHORE : ALEXANDER G. WESMAN

The actual problems, test scores and counseling reports of thirty high school students, with comments on how the cases were handled.

96 pages $1.75 per copy

**Additions to the series of . . .**

# Goldstein-Scheerer Tests

*of Abstract and Concrete Thinking*

| | Test Set | 50 Blanks |
|---|---|---|
| Object Sorting Test | $13.50 | $4.80 |
| Color Sorting Test | 9.50 | 2.70 |

For description, see *Psychological Monographs*, 1941, No. 239, pages 58-109

**Now ready -- 1951 Revision . . .**

# Bellak TAT Blank

Revised Manual, 35 cents Trial Package, 60 cents

Quantity prices remain as given on page 46 of our 1951 Catalog

THE PSYCHOLOGICAL CORPORATION

522 FIFTH AVENUE NEW YORK 18, N. Y.